

AUGMENTED REALITY AND INTRAOPERATIVE C-ARM CONE-BEAM  
COMPUTED TOMOGRAPHY FOR  
IMAGE-GUIDED ROBOTIC SURGERY

by

Wen P. Liu

A dissertation submitted to Johns Hopkins University in conformity with the  
requirements for the degree of Doctor of Philosophy

Baltimore, Maryland

June, 2014

© 2014 Wen P. Liu

All Rights Reserved

# Abstract

Minimally-invasive robotic-assisted surgery is a rapidly-growing alternative to traditionally open and laparoscopic procedures; nevertheless, challenges remain. Standard of care derives surgical strategies from preoperative volumetric data (i.e., computed tomography (CT) and magnetic resonance (MR) images) that benefit from the ability of multiple modalities to delineate different anatomical boundaries. However, preoperative images may not reflect a possibly highly deformed perioperative setup or intraoperative deformation. Additionally, in current clinical practice, the correspondence of preoperative plans to the surgical scene is conducted as a mental exercise; thus, the accuracy of this practice is highly dependent on the surgeon's experience and therefore subject to inconsistencies.

In order to address these fundamental limitations in minimally-invasive robotic surgery, this dissertation combines a high-end robotic C-arm imaging system and a modern robotic surgical platform as an integrated intraoperative image-guided system. We performed deformable registration of preoperative plans to a perioperative cone-beam computed tomography (CBCT), acquired after the patient is positioned for intervention. From the registered surgical plans, we overlaid critical information onto the primary intraoperative visual source, the robotic endoscope, by using augmented reality. Guidance afforded by this system not only uses augmented reality to fuse virtual medical

information, but also provides tool localization and other dynamic intraoperative updated behavior in order to present enhanced depth feedback and information to the surgeon. These techniques in guided robotic surgery required a streamlined approach to creating intuitive and effective human-machine interferences, especially in visualization.

Our software design principles create an inherently information-driven modular architecture incorporating robotics and intraoperative imaging through augmented reality. The system's performance is evaluated using phantoms and preclinical *in-vivo* experiments for multiple applications, including transoral robotic surgery, robot-assisted thoracic interventions, and cocheostomy for cochlear implantation. The resulting functionality, proposed architecture, and implemented methodologies can be further generalized to other C-arm-based image guidance for additional extensions in robotic surgery.

**Advisor**

Russell H. Taylor, Ph.D.

John C. Malone Professor of Computer Science, Johns Hopkins University

**Co-Advisors**

Peter Kazanzides, Ph.D.

Associate Research Professor of Computer Science, Johns Hopkins University

Jeremy D. Richmon, M.D.

Associate Professor of Otolaryngology-Head and Neck Surgery, Johns Hopkins Hospital

Jonathan M. Sorger, Ph.D.

Director of Medical Research, Intuitive Surgical Inc., Sunnyvale, CA



# Acknowledgement

I feel genuinely privileged to recognize and acknowledge the many wonderful individuals whose positivity has propelled me through the pursuit of advanced studies. First, I need to thank my Ph.D. advisors Drs. Russell Taylor, Peter Kazanzides, Jeremy Richmon and Jonathan Sorger for their unwavering belief in my research interests and me. Their expertise and wisdom is second-to-none in this field, but it is their patience, collaborate spirit and generous mentorship that have guided me through such amazing opportunities to learn and cultivate a passion for medical robotics. The multidisciplinary aspect of this field (i.e., what drew me back to graduate school), requires an extension of thanks to Drs. Mehran Armand, Mahdi Azizian, Kevin Cleary, Gary Gallia, Greg Hager, Masaru Ishii, Iulian Iorchata, Yoshito Otake, Douglas Reh, Ken Salisbury, Jeff Siewerdsen; they have all been wonderful principal investigators and research advisors in my additional projects at Johns Hopkins University and other institutes.

Without generosity of funding sources, these pursuits of research would not be possible; in this regard, I would like to acknowledge Intuitive Surgical, Inc.; NIH; the Swirnow Family Foundation; the Chertkof Fellowship; the Department of Otolaryngology – Head and Neck Surgery, Johns Hopkins Hospital; NASA; and Hansen Medical, Inc.

I am profoundly in debt to the gifted engineers and academic consultants that LCSR is fortunate to retain, especially Anton Deguet and Balazs Vagvolgyi, for their invaluable technical solutions, time, and advice.

Fellow student colleagues at CISST-ERC/LCSR and other labs at Hopkins have provided incredible support through amazing friendships incubated during my time in graduate school. They include Paul, Ja, Sohrab, Eric, Blake, Dan, Adam, Ali, Lukas, Orhan, Matt, Issel, Carlos, Vi, Harmony, Alexis, Jon, Tomo, Tricia, Sue Sue, Ioana, Tian, Marcin, Omar, Julian, Nathan, Min Yang, Tamas, Seth, Yixin, Xingchi, Eric, Tze-Yuan, Yan Yan, Topher, Amy, Shahin, Hongliang, Carol, Kevin, Keith, Kel, Manu, Ryan, Hari, Chris, Nishikant, Zihan, Anand, Tutkun, Ehsan, Simon, my best friend Lye-Lin.

Special thanks to the Lin, Cheng and Wang families for your friendship and care since childhood. To the Chuang-Deglow-Lin clan: Shu-yen, Kyle, Kelly, and Kai-wen, you have made me feel like family.

To my family, extended in Taiwan and immediate in the US, my dad, sister and mom, who has passed on, I express my thanks, gratitude, and love.

# Dedication

*To my family, especially my mom, who valued education above  
all.*

# Contents

List of Figures .....	xv
List of Tables .....	xxi
Common Abbreviations .....	xxiii
1 Introduction.....	1
1.1 Background.....	2
1.2 Summary of Approach.....	3
1.3 Clinical Motivations .....	5
1.3.1 Oropharyngeal Cancer .....	5
1.3.2 Robot-assisted Thoracic Surgery .....	7
1.3.3 Skull Base Lesions .....	9
1.3.4 Robot-assisted Cochleostomy for Cochlear Implants.....	10
1.4 Research Problem Statement .....	13
1.5 Technical Barriers.....	15
1.6 Contributions .....	17
2 Related Work .....	20
2.1 Computer-Integrated Surgery .....	20

2.2	Computer-Integrated Surgery Paradigms: Surgical Computer-Assisted Design (CAD)/Computer-Assisted Manufacturing (CAM) and Surgical Assistance.....	21
2.2.1	Surgical CAD/CAM.....	22
2.2.2	Surgical Assistance .....	25
2.2.2.1	<i>da Vinci</i> Robotic System.....	27
2.2.2.2	Mixed Paradigm Systems in Minimally-Invasive Surgery .....	29
2.2.3	Chapter Summary and Discussion .....	31
3	A Modular System for Image-Guided Robotic Surgery .....	34
3.1	Clinical Workflow .....	34
3.2	System Requirements .....	38
3.2.1	Functional Requirements .....	38
3.2.2	Performance Requirements.....	40
3.2.2.1	Accuracy .....	40
3.2.2.2	Visualization .....	42
3.3	Architecture: A Modular Design .....	44
3.3.1	Visualization Module.....	51
3.3.2	User Interface Module .....	52
3.3.3	Robotic/Navigation Module.....	55
3.3.3.1	<i>da Vinci S</i> .....	56
3.3.3.2	<i>da Vinci Si</i> .....	57
3.3.3.3	<i>da Vinci Sp</i> .....	57

3.3.3.4	Medtronic Stealthstation .....	59
3.4	Imaging System: C-arm CBCT, Angiography, and Fluoroscopy .....	60
3.5	Calibration .....	63
3.5.1	Optical Camera Calibration .....	63
3.5.2	Hand-Eye Calibration .....	64
3.5.3	C-arm Calibration .....	66
3.6	Registration .....	68
3.6.1	Deformable CT to CBCT Registration .....	68
3.6.2	Video to CBCT Rigid Registration .....	70
3.6.3	Video to Critical Structures Vision-Based Rigid Registration .....	72
3.7	Chapter Summary and Future Work .....	73
3.8	Recapitulation of Contributions .....	74
4	Integration of C-arm Fluoroscopy for Intraoperative Image-Guided Robotic Surgery .....	77
4.1	Conventional Navigation: Optical and Electromagnetic Tracking Systems .....	78
4.2	C-arm Fluoroscopy-Based Navigation .....	81
4.3	C-arm Fluoroscopy for Intraoperative Image-Guided Robotic Surgery .....	84
4.3.1	C-arm and Robotic Workspace Analysis .....	84
4.3.1.1	<i>da Vinci S/Si</i> .....	85
4.3.1.2	<i>da Vinci Sp</i> .....	91
4.3.2	Digital Tomosynthesis .....	93

4.3.3	Intraoperative 3D Localization from Two X-ray C-arm Views .....	95
4.3.4	Experimental Validation of 3D Localization from Intraoperative X-ray Fluoroscopy for Robot-Assisted Thoracic Surgery .....	99
4.3.4.1	<i>Thoracic Phantom Image Acquisition</i> .....	100
4.3.4.2	<i>Image Analysis for Dual X-ray Projection Geometries</i> .....	100
4.3.4.3	Discussion .....	110
4.4	Chapter Summary and Future Work .....	113
4.5	Recapitulation of Contributions .....	115
5	Augmented Reality for Image-Guided Robotic Surgery .....	117
5.1	Related Work .....	118
5.2	Stereoscopic Video Augmentation .....	123
5.2.1	<i>da Vinci S/Si</i> .....	124
5.2.2	<i>da Vinci Sp</i> .....	128
5.3	Tool Tracking .....	131
5.4	Enhanced Depth Perception .....	133
5.4.1	Dynamic Augmentation .....	134
5.4.2	Virtual Perspective .....	135
5.5	Preclinical Experiments: Evaluation of Augmented Reality for Transoral Robotic Surgery .....	137
5.5.1	TORS Phantoms .....	137
5.5.2	Robotic Experimental Protocol .....	142

5.5.3	Results.....	145
5.6	Chapter Summary and Future Work.....	151
5.7	Recapitulation of Contributions.....	153
6	Preclinical Studies: Application Validation.....	155
6.1	Transoral Robotic Surgery.....	156
6.1.1	Deformable Image Registration for Cone-Beam CT Guided Transoral Robotic Base-of-Tongue Surgery .....	156
6.1.1.1	Introduction.....	157
6.1.1.2	Materials and Methods.....	157
6.1.1.3	Results.....	159
6.1.2	Intraoperative Cone-Beam CT Guidance for Transoral Robotic Surgery ..	161
6.1.2.1	Introduction.....	161
6.1.2.2	Materials and Methods.....	162
6.1.2.2.1	System Overview and Workflow.....	162
6.1.2.2.2	Porcine Models .....	163
6.1.2.2.3	Image Guidance .....	165
6.1.2.3	Experiments .....	169
6.1.2.4	Results.....	170
6.1.2.5	Discussion.....	173
6.2	Cardiothoracic Robotic Surgery .....	174
6.2.1	A Pilot Study of Augmented Reality from Intraoperative CBCT for Image- Guided Thoracic Robotic Surgery .....	175



6.2.1.1	Introduction.....	175
6.2.1.2	Material and Methods .....	176
6.2.1.3	Results.....	179
6.2.2	Discussion .....	180
6.3	Endoscopic Skull Base Surgery.....	181
6.3.1	A Clinical Pilot Study of a Modular Video-CT Augmentation System for Image-Guided Skull Base Surgery.....	181
6.3.1.1	Introduction.....	182
6.3.1.2	System Architecture.....	182
6.3.1.3	Surgical Planning.....	183
6.3.1.4	Video-CT Registration.....	184
6.3.1.4.1	Camera Calibration .....	184
6.3.1.4.2	Hand-eye Calibration .....	184
6.3.1.5	Preclinical Tests: Cadaver Studies.....	185
6.3.1.6	Clinical IRB Case Study .....	187
6.3.2	Discussion .....	190
6.4	Robotically-Assisted Cochlear Implant.....	191
6.4.1	Cadaveric Feasibility Study of da Vinci Si-Assisted Cochlear Implant With Augmented Visual Navigation for Otologic Surgery .....	191
6.4.1.1	Introduction.....	192
6.4.1.2	Methods.....	193
6.4.1.3	Results.....	197
6.4.2	Discussion .....	200

6.5	Chapter Summary and Future Work .....	205
7	Summary and Conclusions .....	209
7.1	Future Improvements .....	210
7.1.1	Optimizing Modules and Functionalities .....	210
7.1.1.1	Augmented Reality .....	211
7.1.1.2	Improved Tool Tracking .....	211
7.1.1.3	Updates for Non-Rigid (Deformable) Tissue Deformation .....	212
7.2	Future Research Directions .....	212
7.2.1	Multi-Modality Integration .....	213
7.2.2	Continuous Assessment in a Closed Loop System .....	213
7.3	Summary .....	214
	Bibliography .....	218
	Vita .....	234
	Curriculum Vitae .....	236

# List of Figures

Figure 1.1 Diagram showing parts of the oropharynx <sup>26</sup> .....	7
Figure 1.2 Diagram showing parts of the skull base from accessible with an endonasal approach <sup>33</sup> .....	10
Figure 1.3 Diagram showing placement of the components of a cochlear implant with respect to the auditory neuropathology system <sup>46</sup> .....	12
Figure 3.1 Image guidance workflow of a mixed paradigm system, featuring TORS as the main clinical motivation that requires deformable registration of surgical planning from preoperative CT to robotic stereoscopic video using intraoperative CBCT.....	37
Figure 3.2 A diagram of the generic modular architecture of the proposed intraoperative image-guided robotic surgical system.....	45
Figure 3.3 An architectural component diagram detailing the libraries supporting stereo video augmentation and modular design of the proposed intraoperative image-guided robotic system used for TORS. ....	48
Figure 3.4 An architectural component diagram detailing the classes supporting video augmentation and emphasizing the modular design of the system architecture. The system is an extension of the TREK architecture for image-guided surgery, binding <i>cisst</i> /SAW libraries for real-time tracking and registration with 3D Slicer libraries for front-end visualization. The specific embodiment described in this dissertation was intended to streamline calibration processes in a manner suitable to clinical use by a trained OR technologist without disruption of OR workflow.....	50

Figure 3.5 Visualization and 3D user interface for transoral robotic surgery (synthetic head phantom).....	54
Figure 3.6 Visualization and user interface for endoscopic endonasal skull base surgery using TREK and Medtronic StealthStation.....	55
Figure 3.7 Labeled components of the research system, <i>da Vinci Sp</i> .....	59
Figure 3.8 Artis zeego CBCT angiography of a porcine liver with a synthetic tumor (white in triplanar views), arterial (red), and venus portal (blue) phases. ....	63
Figure 4.1 Photographs of the <i>da Vinci Si</i> -zeego workspace configuration for transoral robotic surgery. ....	89
Figure 4.2 Photographs of the <i>da Vinci Si</i> -zeego workspace configuration for cochlear implant . ....	90
Figure 4.3 Photographs of the <i>da Vinci</i> -OEC workspace configuration (a) Position of PSC at $\sim 30^\circ$ to table (b). OEC 9600 at end of scan limit ( $30^\circ$ ) (c). Placement of <i>da Vinci</i> robotic arms for thoracic intervention.....	90
Figure 4.4 Photographs of the <i>da Vinci Si</i> -zeego workspace configuration for thoracic intervention (a) zeego positioned at table head with docked patient-side cart. zeego at the start (b) and (c) end of a $\sim 45^\circ$ * scan range. ....	91
Figure 4.5 Photographs of the <i>da Vinci Sp</i> -zeego workspace configuration for gynecology. ....	92
Figure 4.6 Photographs of the <i>da Vinci Si</i> -zeego workspace configuration for hepatic intervention. ....	93
Figure 4.7 Single axial slice of canine thoracic phantom from digital tomosynthesis of (a) $30^\circ$ (b) $60^\circ$ (c) $90^\circ$ (d) $180^\circ$ scans.....	95
Figure 4.8 Radiographs during 2D-3D registration and DRR for ground truth assessment. (a) Anterior-posterior radiograph with gradient (red) of registered DRR. (b) Single coronal slice of reconstructed canine thorax with segmented heart (orange), mock tumor (yellow) and numbered peri-tumor fiducials (blue). (c) Sagittal and (d) coronal fluoroscopic image with forward projections of segmented 3D targets. ...	103

Figure 4.9 Results from dual projection experiments focused on Artis zeego extrinsic uncertainties. Table (left, top) TRE from perturbing fluoroscopic data. Table (left, middle). TRE after 2D-3D registration on perturbed DRR data. Table (left, bottom) TRE after after 2D-3D registration on perturbed fluoroscopic data. (a). Box plot of the zeego extrinsic uncertainty on DRR TRE (experiment #4). (b). Box plot of the zeego extrinsic uncertainty on fluoroscopic TRE (experiment #3). .....	107
Figure 4.10 (Top: Experiment #5) Box plot of the effect of angular difference on TRE with various rotational perturbations. (Bottom: Experiment #6) Box plot of the effect of angular difference on TRE with various translational perturbations. ....	109
Figure 4.11 (Top: Experiment #7) Box plot of the effect of angular difference on TRE with various focal length perturbations. (Bottom: Experiment #8) Box plot of the effect of angular difference on TRE with various optical center perturbations. ....	110
Figure 5.1 Experimental setup using a <i>da Vinci S</i> system with stereoscopic video overlay of spherical soft-tissue targets (magenta) segmented from intraoperative CBCT..	127
Figure 5.2 A diagram showing the hierarchy of coordinate systems maintained by the 3D user interface. ....	127
Figure 5.3 A photograph of the patient side cart of the <i>daVinci Sp</i> , showing the entry guidance manipulator (i.e., the main structural chassis). ....	130
Figure 5.4 A photograph of the <i>da Vinci Sp</i> , a single-port robotic system with 4 channels within a 25 mm trocar that accommodates a flexible endoscope and 3 articulated instruments. ....	130
Figure 5.5 Photograph of the feature-based markers attached to a 5 mm monopolar cautery instrument during vision-based tool tracking. ....	132
Figure 5.6 Screenshot of the enhanced augmented reality scene with a virtual, orthogonal perspective in the lower left corner. As the tracked needle tip approaches the embedded target, enhancements include the change in color (red) of the sphere, a distance label (white on end effectors), and the virtual perspective (picture-in-picture) showing a red avatar of the tracked point inside the target (yellow). ....	136

Figure 5.7 Photographs and CBCT images of a porcine tongue positioned by (a,b) a flat, preoperative template and (c,d) a curved, extended intraoperative position. The bright punctuated lesions in the scans represent the Teflon targets. ....	139
Figure 5.8 Cadaver head positioned in (a) a preoperative pose and (b) imaged in CT. The same cadaver was positioned in (c) an intraoperative pose [tongue sutured and extended] (d) and imaged using C-arm CBCT. ....	139
Figure 5.9 Cadaver setup emulating the intraoperative setup for CBCT-guided TORS. (a) Prototype mobile C-arm for intraoperative CBCT. Cadaver mounted in a CT-compatible frame using nylon strings to retract the jaw and tongue as shown in (b) and (c). ....	140
Figure 5.10 Simulated current practice in which the surgeon has access to only the (a) raw endoscopic image or (b) preoperative CT to help guide the surgical navigation. Our proposed image guidance workflow showing (c) a 3D image overlay of the targets on the tongue derived from the (d) preoperative plan registered to intraoperative CBCT. ....	144
Figure 5.11 Sagittal slice of post experiment CBCT from the Model C phantom (cadaver head). Target Localization Error (TLE) is deconstructed into 4 types: TLE (Edge, Center, Projection, and Depth). ....	148
Figure 5.12 Clockwise from upper left, box plots of TLE for Center, Edge, Depth, Projection for porcine experiments. ....	151
Figure 5.13 From left to right, box plots of TLE for Center, Edge, Projection, Depth for cadaveric experiments. ....	151
Figure 6.1 Images and segmentation of a cadaver specimen in (a) Preoperative CT (I0) and (b) Corresponding CBCT (I1) in the intraoperative state (mouth open and tongue retracted), with segmentation mask in red. ....	158
Figure 6.2 The transformed cadaveric tongue and TRE following each step of the registration framework: (a) GM rigid; (c) GM nonrigid; and (b) Demons. ....	160

Figure 6.3 (a) Single axial slice from CBCT of an <i>ex vivo</i> pig tongue phantom with embedded tumor (green). (b) Single sagittal slice CBCT angiography of an <i>in vivo</i> pig phantom with segmented models of the right lingual artery (orange), and two base-of-tongue tumors (right in yellow, left in blue). (c) Photograph of an <i>ex vivo</i> pig tongue phantom affixed with green registration fiducials. (d) Photograph of an <i>in vivo</i> pig phantom supine and readied for tumor placement. ....	165
Figure 6.4 Screen capture of an <i>ex vivo</i> phantom experiment using video augmentation of margins (green sphere) and tool tracking in novel views (lower left picture-in-picture) for image guidance. ....	166
Figure 6.5 Screen capture of lingual dissection during an <i>in vivo</i> porcine lab experiment using video augmentation as image guidance. ....	168
Figure 6.6 Resected tumor and margins with corresponding slices (tumor in high intensity, white) and segmented volumes (tumor in blue, intersection with ideal margin in yellow) from postoperative CBCT. ....	172
Figure 6.7 (a) Axial and (b) coronal slice of lateral inflated lung. (c) Axial and (d) coronal slice of lateral collapsed lung with a medial mock tumor indicated by red arrow. ....	177
Figure 6.8 Overlays from anatomic registration of (a) pulmonary landmarks and (b) mock tumor (c) Segmentation of critical structures in CBCT (sagittal slice of lateral, inflated data). (d) Wedge resection. ....	179
Figure 6.9 Screenshot of the video-CT system's interface including the augmented display and triplanar view of the CT and surgical plan. ....	184
Figure 6.10 Video-CT Augmentation workflow for image-guided endoscopic endonasal surgery.....	186
Figure 6.11 (a) Example planning data shown in a sagittal CT slice of a cadaver employed in preclinical evaluation. Target structures include the anterior skull base (pink), the inferior clivus (green), the superior clivus (blue), and the surgical target (pituitary, in red). The clinical pilot study involves side-by-side display of (b) conventional (non-augmented) endoscopic video and (c) the experimental system for augmentation and real-time overlay of registered planning data within the video	

scene. The carotid arteries, which in this case have a narrow inter-carotid distance, are overlaid in red in a trans-sphenoid clival drillout procedure. ....	187
Figure 6.12 (a)-(c) Three perspectives of video-CT overlay as the endoscope is panned left-to-right and anterior-to-posterior in a preclinical evaluation following a cadaveric clival drillout. The images show overlay of the carotid arteries (red), optic nerves (blue), and pituitary gland (purple) in trans-sphenoid approach to the skull base. ....	187
Figure 6.13 Clinical study operating room setup, showing components from the standard-of-care and video augmentation system. ....	190
Figure 6.14 Layout of the operating room with the <i>da Vinci Si</i> for Case 1. Inset is a close-up of the initial position of the endoscope, suction/irrigator, and drill attached with the custom tool adapter. ....	196
Figure 6.15 (a) Master-slave-assisted mastoidectomy and cochleostomy on cadaveric right temporal bone completed with augmented reality, registered using three fiducials. (b) Coronal slice in preoperative cone-beam computed tomography showing segmentation of the critical structures. (c) Monocular screen capture of the right eye during cochleostomy with video augmentation of the segmented models. ....	197
Figure 6.16 Axial slice from postoperative CBCT showing the successful placement of a phantom implant wire (yellow) in the cochlea. ....	200



# List of Tables

Table 3.1 Research Problems (Section 1.4) with Contributions to Overcome Technical Barriers (Section 1.5) .....	74
Table 4.1 Experimental Protocol for 3D Localization Experiments.....	103
Table 4.2 OEC 9600 Extrinsic Experimental Results.....	106
Table 4.3 Summary of Workspace Analysis.....	115
Table 4.4 Research Problems (Section 1.4) with Contributions to Overcome Technical Barriers (Section 1.5) .....	115
Table 5.1 Experiments for Phantom Target Localization .....	147
Table 5.2 Results from Phantom Target Localization .....	148
Table 5.3 Statistical significance (p-values) in the measured differences in TLE between the various modes of operation in the porcine phantom experiments. Statistically significant results (p-value < 0.05) are in <u>underlined bold</u> . .....	149
Table 5.4 Statistical significance (p-values < 0.05) were achieved in measurements of all TLEs in the cadaver specimen between the conventional mode of operation (C1) and the proposed workflow (C2) integrating intraoperative imaging and endoscopic overlay.....	150
Table 5.5 Research Problems (Section 1.4) with Technical Barriers Resolved (Section 1.5) .....	153
Table 6.1 TORS Experiments .....	170

Table 6.2 TORS Experimental Results .....	172
Table 6.3 Summary Questionare For Expert Assessment of Video-CT in Skull Base Surgery .....	189
Table 6.4 Research Problems (Section 1.4) with Technical Barriers Resolved (Section 1.5) through Inline Contributions.....	207

# Common Abbreviations

CT	Computed Tomography
MR	Magnetic Resonance
CBCT	Cone-Beam Computed Tomography
TORS	Transoral Robotic Surgery
US	Ultrasound
MIS	Minimally-Invasive Surgery
OP	Oropharyngeal
SSCa	Squamous Cell Carcinoma
FDA	U.S. Food Drug Administration
CI	Cochlear Implant
VATS	Video-Assisted Thoracoscopic Surgery
CTA	Computed Tomography Angiography
CIS	Computer-integrated Surgery
API	Application Programming Interface
CAD	Computer-Assisted Design
CAM	Computer-Assisted Manufacturing
PUMA	Programmable Universal Machine for Assembly
EM	Electromagnetic
THA	Total Hip Arthroplasty
IDE	Investigational Device Exemption
SAGES	Society of American Gastrointestinal and Endoscopic Surgeon
NASA	National Aeronautics and Space Administration
SRI	Stanford Research Institute
RCM	Remote Center of Motion
SSC	Surgeon-side Console
AR	Augmented Reality

IGRS	Image-guided Robotic Surgery
PSC	Patient-Side Cart
VTK	Visualization Toolkit
UIL	User Interface Library
MTM	Master Manipulator
PSM	Patient-Side Manipulator
II	Image Intensifier
3DUI	3D User Interface
TRE	Target Registration Error



# 1 Introduction

Robotic surgery requires intraoperative precision and extensive understanding of the three-dimensional topography of the surgical target, the spatial relationship of surrounding vasculature, and the vital anatomy with respect to instrument positioning. Multimodal imaging, including high-resolution computed tomography, angiography, magnetic resonance imaging, and ultrasound, provides visualization of significant critical anatomical boundaries used in standard of care for diagnostics and preoperative planning.<sup>1-3</sup> Clinicians derive surgical strategy from such volumetric data, which include planned traversals to resection targets and controlling or preserving critical functional structures.

The emergence of minimally-invasive surgery (MIS) has accentuated a gap that currently exists between the display of preoperative surgical plans and the camera of video-based (i.e., endoscopic) interventions. During video-based and robot-assisted interventions, preoperative images are generally viewed on an external monitor; therefore, these images are not aligned with the patient anatomy. Although the magnification and stereo capabilities of current endoscopes provide unparalleled high-resolution images of the surgical field during robot-assisted surgery, tactile feedback is diminished or absent, which increases the surgeon's reliance on visual cues. As such, the novice clinician may need to rely mostly on his knowledge of anatomic structures, which may be insufficient to anticipate surgical strategy beyond the endoscopic view. A loss of orientation or localization increases the inherent risks of breaching critical structures,

especially in the hands of inexperienced surgeons.

## 1.1 Background

Minimally-invasive surgeries (MIS) visualize the operative workspace with optical cameras (i.e., laparoscopic cameras, thoroscopic cameras, and endoscopes) inserted through small incisions or natural lumens/orifices. Compared to traditional open surgery, MIS has advantages<sup>4-9</sup> that directly impact patient quality of life issues, including: less estimated blood loss and transfusions, fewer complications, lower mortality rate, and shorter length of hospital stay. However, minimally-invasive procedures create a challenging environment in which information feedback is image-based, rather than directly tactile, thereby adding increased complexity. During the last decade, robotics and image guidance systems have been introduced to operating rooms to overcome some of the current challenges.

Robotic or robot-assisted minimally-invasive surgery has been used in a variety of surgical specialties, including neurosurgery, head and neck surgery, orthopedics, cardiothoracic surgery and urology, among others<sup>10</sup>. Systems such as the *Zeus* (formerly ComputerMotion Inc., Goleta, CA, USA) and *da Vinci* (Intuitive Surgical, Sunnyvale, CA, USA) use a tele-robotic approach that interprets input motion from a surgeon to direct dexterous end effectors. By filtering tremor, magnifying high definition stereoscopy, and maintaining a mechanical fulcrum at the incision sites, these systems help reduce the complex mental to physical coordination needed in a constrained workspace, especially when compared to traditional laparoscopic approaches.

Concurrently, modern multi-modal diagnostic imaging continues to evolve towards high resolution digital 3D volumes that target specific physiology and function. For example, computerized tomographic angiographies can map vascular structures with applications in treating pulmonary embolism, carotid/vertebral dissection, and locating aneurysms<sup>11</sup>. Image guidance has been integrated with a subset of medical robotic interventions to improve upon standard of care for neurosurgery<sup>12,13</sup>, retinal surgery<sup>14,15</sup>, and orthopedics<sup>16,17</sup>. For example, *ROBODOC*<sup>18</sup> (Curexo Technology Corp, Fremont, CA) and *MAKO*<sup>19</sup> (MAKO Surgical Corp, Ft. Lauderdale, Florida) use preoperative CT to create a surgical plan for total knee and hip arthroplasty.

The orthopedic and neurosurgical disciplines have mainly adopted guidance from preoperative imaging because they can assume a rigid anatomy, and the surgeons can directly see and touch the operative workspace. However, for many other video-based interventions, particularly operating with soft tissue targets, standard practice involves displaying 3D medical data (CT/MR) in 2D displays, separately from the endoscope. The physician needs to identify anatomical structures in videography and mentally establish the spatial relationships from preoperative plans. The ability to continuously maintain correspondence creates a steep learning curve, remains a function of experience, and therefore is subject to inconsistencies. Patient positioning and intraoperative deformation from surgical intervention (i.e., retraction, dissection, resection) further compounds the problem, altering the anatomy even further from preoperative image acquisitions.

## **1.2 Summary of Approach**

The objective of this work was to develop a versatile image guidance system for video-based minimally-invasive surgery. We integrated a Siemens Artis zeego, a high-



end robotic C-arm system, with the *da Vinci* platform, a state-of-the-art clinical robotic system, by extending existing frameworks and libraries in our development of a modular architecture capable of supporting multiple surgical applications. From standard preoperative diagnostic image data (e.g., CT, MRI), we identified critical structures (e.g., a tumor, adjacent arteries, and nerves) that were initially registered with perioperative C-arm image data that could be directly overlaid onto intraoperative endoscopic video as image guidance. We posited that such augmentation of relevant anatomy would improve navigation, spatial orientation, confidence, and tissue margins.

The layout describing the proposed work is organized as follows:

The remainder of Chapter 1 presents several clinical applications in otolaryngology, head and neck surgery, thoracic surgery, and neurosurgery that are motivating this work. Chapter 2 gives a high-level historical background of the development of related work in image guidance for robotic surgery. Chapter 3 describes the design of our modular architecture, including performance accuracy, individual components, and their functional services, in addition to component-based provided and required interfaces. Chapter 4 shows a more detailed exploration of advantages given by an interventional C-arm to provide intraoperative updates through X-ray fluoroscopic imaging. Furthermore, Chapter 5 describes our key visualization component, where we explored methods of augmented information feedback with video augmentation, dynamic visual cues, and enhanced depth perception. The introduction of our image-guided surgical system for current surgical interventions required an efficient integration into existing surgical workflows. Thus, in order to demonstrate feasibility and potential advantages of our approach as compared to simulated current practices, we conducted a

validation study for each of our target clinical applications, as detailed in Chapter 6. Chapter 7 presents our conclusions, including immediate next steps, as well as potential future work.

## **1.3 Clinical Motivations**

### **1.3.1 Oropharyngeal Cancer**

The oropharynx is the middle part of the throat, which includes the base of the tongue, the tonsils, the soft palate and the post pharyngeal wall (Figure 1.1). Oropharyngeal (OP) cancer occurs when malignant cells form in the tissue of the oropharynx. Squamous cell carcinomas (SCCa), which constitute the majority of OP cancer<sup>20</sup>, are flat, scale-like cells that normally arise from the lining of the mouth and throat. OP cancer can be generally divided into two types: HPV-positive, which are related to human papilloma virus (HPV) infection, and HPV-negative cancers, which are typically linked to alcohol or tobacco use<sup>21</sup>. The association of HPV with SCCa is a rising national health concern, accounting for 70% of OP cancer in 2000, as compared to 16% in the 1980s. One unfortunate epidemiological consequence of this trend is that oncologists must treat a younger patient population whose long-term outcomes requires more consideration compared to historically older, cigarette users who have developed HPV-negative cancer<sup>22</sup>.

Standard treatments for OP cancer include radiation, chemotherapy, and open surgery. These invasive techniques often leave collateral damage, debilitating speech and swallow function, especially with open surgery, which requires a transcervical approach, possible mandibulotomy, and tracheotomy. However, recent studies show that among

many of the OP cases surgery often offers the greatest chance of cure<sup>23</sup>. In December of 2009, the US Food and Drug Administration (FDA) approved the use of the *da Vinci* surgical robot to perform transoral robotic surgery (TORS) as a minimally invasive alternative. Studies using TORS have shown safety and feasibility with oncologic equivalence to chemo-radiation therapy and improved function when compared to traditional open surgical approaches for OP cancer<sup>24,25</sup>. However, intraoperatively, the risks of lacerating critical tissues, as well as standard methods needed to identify boundaries between normal and cancerous tissues for optimal preservation of function, remain unaddressed.

Resection of tumors via TORS is achieved en bloc. To approach a tumor, often an individualized process, surgeons rely on pre-operative imaging, physical examination, visual feedback, the input of the bedside assistant, and personal experience. Usually, surgery requires dissection in a plane beyond the tumor margin as the tongue curves towards the vallecula with limited visualization of the true depth of dissection. This unfamiliar orientation contributes to the steep learning curve to the resection of deep base of tongue cancers, compounded by the lack of anatomic landmarks and the potential for significant bleeding.

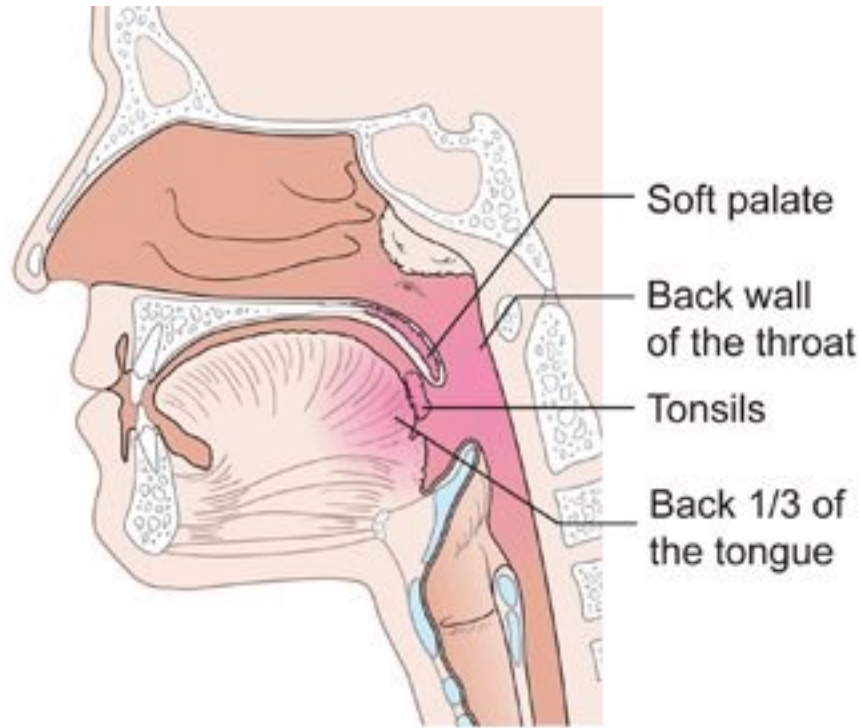


Figure 1.1 Diagram showing parts of the oropharynx<sup>26</sup>.

### 1.3.2 Robot-assisted Thoracic Surgery

Tele-robotic thoracic interventions available with a *da Vinci* system (Intuitive Surgical Inc., Sunnyvale, CA) offer notable advantages in the delicate dissection required with systematic mediastinal or pulmonary lymphadenectomy and other intricate work in the thoracic cavity. Initial results have been promising with respect to improved short-term outcome, as compared to open thoracotomy and even Video-Assisted Thoracoscopic Surgery (VATS)<sup>27</sup>. In fact, several articles have shown the efficacy and safety of robotic pulmonary applications, including lobectomy, segmentectomy, and wedge resections. Recent national<sup>28</sup> and multi-center<sup>29</sup> studies support robotic pulmonary resection as an appropriate alternative to VATS. Furthermore, preliminary results of single institutional

studies on robotic lung segmentectomy<sup>30</sup> support robotic intervention as a feasible and safe approach.

The amount of lung tissue to be resected depends on the etiology and histology of the tumor. Tumor stage is based on the size and/or extent of the primary tumor, whether cancer cells have spread to nearby (regional) lymph nodes, and the level of metastasis, which is the spread of the cancer to other parts of the body. Currently, the standard-of-care for operable lung cancer is a lobectomy with systematic lymphadenectomy. For smaller tumors, however, a less extended anatomical resection, such as a segmentectomy, might become the appropriate surgical treatment in the near future.<sup>31</sup> A pulmonary segment is an anatomical unit that consists of an alveolar duct (bronchus), the air spaces connected with it, their blood vessels (pulmonary artery & vein), lymphatics, and nerves. These structures divide and subdivide further inside the lung parenchyma, making localization, dissection, and resection more difficult. An appropriate anatomical pulmonary resection (e.g., segmentectomy) requires precise knowledge of the relevant pulmonary anatomy, as well as excellent three-dimensional spatial orientation.

Preoperative volumetric data from computed tomography (CT) or positron emission tomographic-CT are the major diagnostic tool used to achieve adequate staging information. These images are acquired with the patient in the supine position. Furthermore, in order to create a workspace for robotic intervention, the lung must be collapsed with the patient rotated 90° laterally and overextended in the coronal plane, presenting a thoracic workspace that is deformed from the workspace in preoperative image acquisition.

### **1.3.3 Skull Base Lesions**

As the most inferior area of the skull, the skull base forms the floor of the cranial cavity and separates the brain from other facial structures. It can be divided into three regions: the anterior, middle, and posterior cranial fossae. In the case of pituitary adenomas, surgeons require access to the central compartment of the skull base, the middle cranial fossa, which contains the pituitary gland. Neurosurgeons and otolaryngologists use minimally invasive techniques to reach and remove tumors from the skull base and intracranial cavity by operating through the nose and paranasal sinuses (Figure 1.2). Abnormalities of the pituitary can affect hormonal imbalance; for example, in the case of Cushing's Disease, pituitary adenomas cause excessive secretion of an adrenocorticotrophic hormone that stimulates the adrenal glands to produce excessive amounts of cortisol. Cushing's Disease may manifest in psychiatric and emotional instability, as well as cognitive difficulties, possibly causing fatigue and weight gain.

Compared to traditional open surgery, endoscopic endonasal and minimally invasive robotic skull base procedures can spare patients from considerable morbidity and complications. These techniques necessitate precise visualization to ensure complete resection within the complex anatomy of the endonasal space<sup>32</sup>. Such skull base pathologies are in close proximity to critical neurovascular structures (e.g., carotid arteries, optical nerves); therefore, injuries can cause significant consequences (e.g., neurological injury and death).

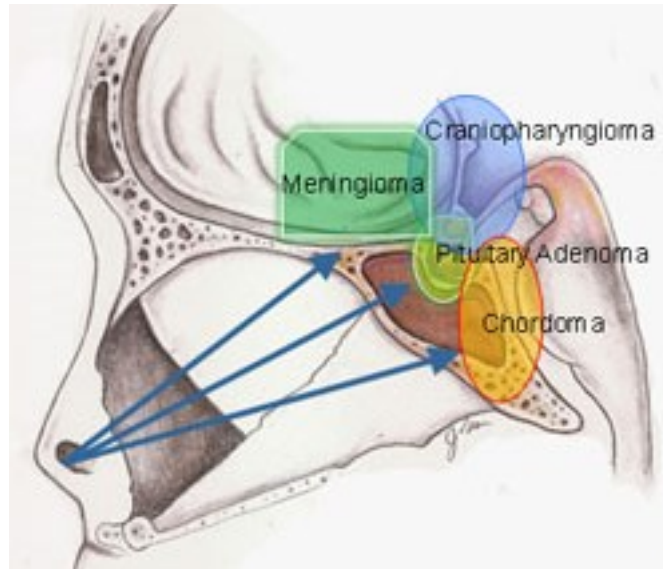


Figure 1.2 Diagram showing parts of the skull base from accessible with an endonasal approach<sup>33</sup>

### 1.3.4 Robot-assisted Cochleostomy for Cochlear Implants

A cochlear implant (CI) is an electronic device approved by the United States Food and Drug Administration that can address both congenital and acquired sensorineural hearing loss. Open surgical techniques to place cochlear implants were pioneered in the 1960s and have been performed in over 200,000 individuals worldwide.<sup>34</sup> The device consists of a microphone, speech processor, transmitter, receiver/stimulator, and an electrode implant array that collects the impulses from the stimulator and sends them to different regions of the auditory nerve. Typically, cochlear implant surgery involves a standardized approach through a mastoidectomy, facial recess drill out, posterior tympanostomy, and then cochleostomy with electrode insertion. Accuracy of both cochleostomy placement and cochlear implant insertion angle has been shown to be

critical for device function and clinical functional outcomes.<sup>35-37</sup> Although there have been many innovations to the implant device itself, particularly with the development of multichannel electrodes, this traditional open surgical approach remains largely unchanged.

Appropriate cochleostomy placement (Figure 1.3) is important to allow for proper cochlear electrode array insertion in the scala tympani, avoiding the scala media and vestibuli. This can be challenging given the depth of the cochlea, which is 30 mm within the temporal bone, and the width of a facial recess, which is usually 2.5 to 3 mm with a “target area” for cochleostomy of 1 mm for appropriate insertion. The current cochleostomy surgical technique requires freehand drilling with a 0.6 to 1 mm diameter burr through the promontory into the scala tympani, optimally avoiding intracochlear trauma. Recent studies, however, suggest that a significant proportion of cochlear implant surgeons do not adequately position the cochleostomy, typically described as inferior or anterior-inferior to the round window<sup>38,39</sup>. Lack of familiarity with the facial recess can result in the surgeon leaving too much undrilled bone overlying the nerve, with an incomplete opening of the facial recess and poor visualization of the round window niche or membrane. Other potential factors that contribute to inadequate cochleostomy placement include variable round window anatomy, a poor angle of visualization approach through a restricted facial recess, and complex inner ear (cochlear/vestibular) malformations.

In contrast to percutaneous robotic methods, and similar to traditional CI, a robot-assisted mastoidectomy and cochleostomy with stereo endoscopy maintains the advantage of possible manual correction through visual cues, allowing the surgeon to



maintain a high level of engagement in intraoperative decision making/dissection and reduces risk of error from over-reliance on a preplanned surgical route. Furthermore, by adopting a robotic CIS approach, clinicians can take advantage of mechanical dexterity from robotic-assistance, as well as image guidance, thereby improving safety and accuracy by including path planning<sup>40</sup> and virtual fixtures<sup>41-45</sup>.

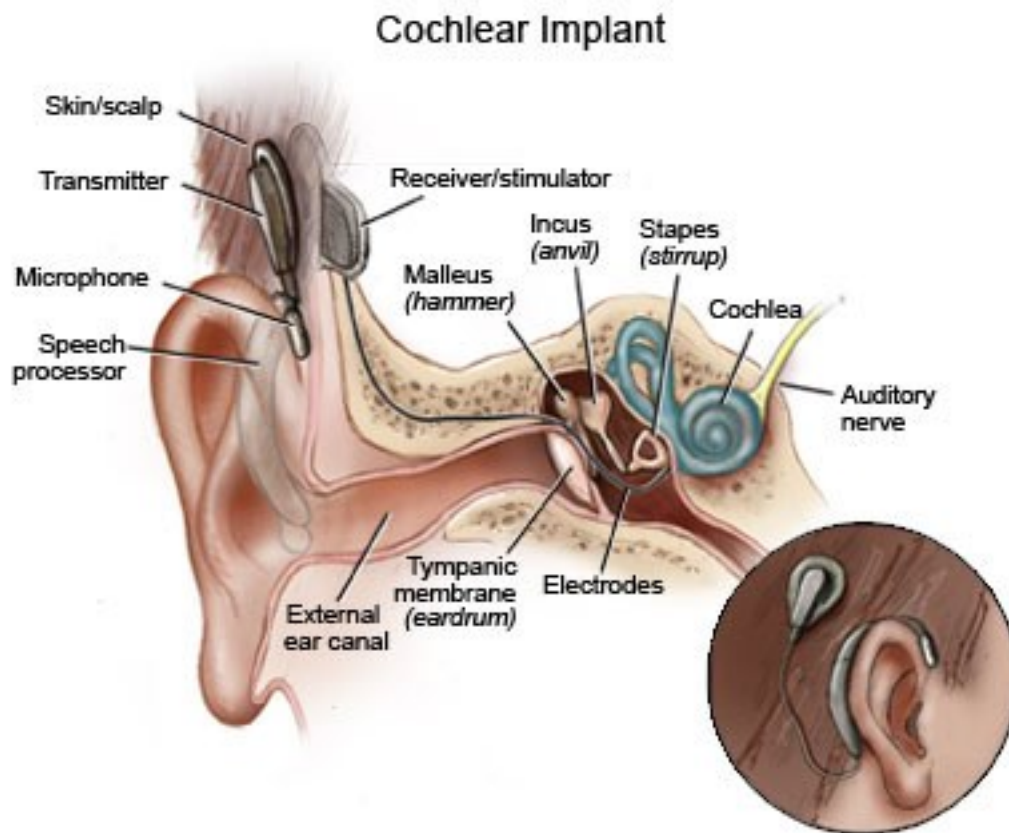


Figure 1.3 Diagram showing placement of the components of a cochlear implant with respect to the auditory neuropathology system<sup>46</sup>.

## 1.4 Research Problem Statement

Oncologic targets can be buried deep in an organ or musculature, presenting a challenge to clearly visualize all but the most superficial aspects of the tumor. During robotic surgery, a lack of haptic feedback in addition to these challenges can lead to potential injury to neurovascular structures, jeopardizing reliable and safe delineation of tumor margins. Clinicians typically use preoperative data<sup>47-50</sup> to plan the surgery by taking advantage of the capabilities of diverse modalities. Perioperatively, this plan needs to be updated with considerations for the possible modifications of a patient's anatomy after preoperative image acquisition; examples of such modifications may include the progress of a disease or pre-interventional treatment such as chemotherapy or radiation therapy, the current filling of the stomach or bladder, and the intraoperative patient position. Intraoperative imaging<sup>51,52</sup> can capture the patient's rigid and deformable movements caused by various factors, such as the patient's cardiac or breathing state and surgical motion. The current gap between volumetric imaging and robotic surgical interventions compels the exploration of methods of visualization for image guidance using preoperative and intraoperative imaging sources. Thus, this work presents the development of a flexible architecture to register and integrate preoperative diagnostic volumetric data with intraoperative C-arm imaging, using augmented reality to provide real-time guidance and tracking for minimally-invasive robotic surgery.

Augmentation using preoperative medical images as guidance has been realized in orthopedic<sup>53</sup>, head and neck<sup>54</sup>, and urologic<sup>49</sup> interventions. In contrast, this system adopts a generic architecture that has been extended and validated for multiple clinical

applications. In addition, augmentation from this system is not only realized with initial rigid registration, but has also been updated to accommodate intraoperative deformation.

- **Problem 1: Mental Correspondence**

In surgical interventions, surgeons remain oriented with respect to critical anatomy, especially after tissue deformation, by mentally mapping volumetric preoperative data to a possibly highly deformed surgical field. Such practice is a function of subjective experience.

- **Problem 2: Perioperative Image Data and Intraoperative Deformation**

Standard preoperative image data sets (i.e., from computed tomography, magnetic resonance imaging) present dense volumetric information that does not always reflect either perioperative setup or intraoperative deformation from interventional motion. By contrast, intraoperative imaging captures not only patient setup, but also anatomical changes since diagnostic acquisitions. However, alignment of preoperative to intraoperative imaging for soft tissue workspaces is not straightforward, requiring non-rigid registration.

- **Problem 3: Effective Image Guidance**

Our proposed solution to implement image guidance by fusing virtual medical information to endoscopic video is a complex sensory experience. This approach

to augmented reality presents challenges in visualization, 3D perception, and user interface. Substantiating the impact of different methods of augmentation, as well as validating the collective effectiveness of the entire system, is a necessary step in order to characterize the usefulness of such emerging technology.

## **1.5 Technical Barriers**

Our proposed approach integrates a C-arm system for intraoperative image guidance in robotic surgery using augmented reality. Here, image guidance aims to complement the physician's ability to understand the spatial structure of the anatomy with respect to preoperative surgical strategy and robotic instrumentation. In this effort, required functionalities include registration between volumetric image data and registration of image data to video, as well as real-time updates of these correlations to reflect surgical motion. Additionally, visualization techniques in real-time stereoscopic video augmentation, feedback of tool information, and explicit 3D depth perception in augmented reality must be engineered. After initial surgical resection, we explore opportunities for intraoperative updates from C-arm and video images. Concepts in computer vision have been applied to these intraoperative images in order to provide real-time updates for registration and tool tracking in a non-rigid environment. Phantoms, *ex vivo*, and *in vivo* models are used to validate feasibility and demonstrate potential impact.

Key technologies enable the deployment of an image-guided system within a feasible clinical workflow. Assuming an idealized high-level timeline view for surgical procedures to include preoperative, intraoperative setup, and update steps, we identify the following technologies associated with different phases: (1) preoperative diagnostic imaging and planning; (2) intraoperative interventional imaging and image processing;

(3) registration; (4) augmented reality; (5) intraoperative tool tracking and updates. More detailed discussion about the requirements of each phase and the development efforts involved can be found in Section 3.1.

The list below summarizes technical barriers that must be overcome in order to address the associated research problems:

- **Problem 1: Mental Correspondence**

- Effective data representation from volumetric CT/MRI/CBCT
- Effective information delivery from volumetric CT/MRI/CBCT

- **Problem 2: Preoperative Data and Perioperative/Intraoperative Deformation**

- Feasible perioperative/intraoperative systems workspace configuration for image acquisition
- Registration of preoperative image data to perioperative/intraoperative image data
- Intraoperative surgical motion

- **Problem 3: Effective Image Guidance**

- Adaptable software architecture and components
- Efficient interfaces and functionality
- Effective systems engineering with evaluation and validation

## **1.6 Contributions**

This dissertation presents a novel paradigm for a modular architecture using augmented reality for C-arm-guided robotic surgery. We address fundamental limitations in minimally-invasive robotic surgery by combining the surgical assistant capabilities of a modern surgical robotic platform with guidance from information afforded by high-end robotic C-arm imaging. The existing hardware from the integrated systems can be leveraged and extended to create novel intuitive and relevant visualization, through human-machine interferences with streamlined design principles. Component-based software design principles are used to build upon open source frameworks and libraries in order to develop a generalized architecture that can be specifically configured for applications to multiple clinical scenarios. The major research contributions realized in this effort are summarized here:

### **Modular System Architecture**

A versatile system using augmented reality and intraoperative imaging for image-guided robotic surgery is implemented and validated. Though more comprehensively tested for our exemplar transoral application, the architecture (Section 3) was conscientiously designed to be modular and versatile, and therefore adaptable to clinical workflows for additional surgical interventions. Different clinical applications may require a different image modality, setup, camera (formats, resolutions, control unit), application programming interface (API), as well as robotic interfaces and tools.

Extendibility to accommodate workflow adaptations as well as validation for each scenario is presented (Section 6).

### **Intraoperative C-arm Integration**

We propose the use of a robotic CBCT-capable C-arm system, such as the Artis zeego (Siemens, Inc.), to capture initial perioperative setup. The acquired CBCT dataset provides an anchor to deformably register standard diagnostic and staging image data. Furthermore, the C-arm can be configured and registered to provide live 2D X-rays and additional 3D scans during a minimally invasive robotic intervention. We present the integration of an intraoperative C-arm with a minimally-invasive robotic system (Section 4), including the analysis of their workspace limitations when configured together for different interventions. For image guidance updates, registration, and tool tracking, we examined methods for volumetric (3D to 3D) registration (Section 3.6.1) and X-ray-based 3D localization from two C-arm views (Section 4.3.3).

### **Augmented Reality for Image Guidance**

Fusion of virtual information, namely the graphical user interface as well as physical interfaces for image guidance in minimally-invasive surgery, can be achieved through many different approaches. The proposed user interface allows surgeons to customize and interact with the virtual environment directly through the current standard surgical console (i.e., a familiar interface). Projective overlay of 3D meshes provides localization of targets within the camera image plane. However, depth perception using

stereoscopic cameras on the *da Vinci* surgeon console is a function not only of camera parameters, but also of the surgeon's natural stereopsis and interpretation of visual cues. In addition to overlays of anatomical targets, we demonstrate novel methods of feedback to communicate distances that are beyond the camera plane and to further highlight additional enhanced 3D depth perception (Section 5.4)

### **Intraoperative Updates**

Superimposed virtual structures from preoperative data provide adequate guidance on approach. However, as dissection and resection progresses, intraoperative updates are required to account for motions that have occurred. We have designed a custom fiducial marker and present a computer vision-based approach (Section 3.6.3) to track intraoperative motion. Furthermore, tool tracking has been improved with forward kinematics using instrument joint encoders, as provided by the API. We correct for an initial offset and derive setup joint corrections through vision-based processing of markers attached to the shaft of the instrument (a proprietary function developed by Intuitive Surgical Inc.). Using these improvements, our method of proposed image guidance includes intraoperative tool tracking (Section 5.3) and rendering of relative information with respect to virtual critical structures.



## 2 Related Work

### 2.1 Computer-Integrated Surgery

Computer-integrated surgery (CIS) systems are designed to enhance the capabilities of a surgeon by managing the flow of *medical information* and to assist in operative *action*. General medical information about human anatomy and variability can be composed from atlases built from a similar population of patients that have undergone the same procedure; however, patient-specific information, derived from volumetric imaging, is preferred. To use computer-integrated surgical systems, clinicians combine specific and general information, alongside their experience, to create a digital model of the individual patient, which is then used to plan surgical strategies.

In the operative action aspect, medical robotics has been developed to provide precision and mechanical dexterity. In this sense, the robot itself is simply an element within a comprehensive computer-integrated surgical system. Generally, robotic devices can be classified into three broad categories: active, semi-active, and passive. An active robot performs parts of the procedure autonomously (e.g., ROBODOC<sup>18</sup>, further discussed in Section 2.2.1). A semi-active robot performs the procedure under the direct control of the surgeon (e.g., Acrobot<sup>55</sup>, Steady Hand Robot<sup>56</sup>). Lastly, a passive robot does not actively perform any part of the procedure, which is fully controlled by the surgeon. Passive robotic devices are often used to place or hold instruments (e.g.,

NeuroMate<sup>57</sup>). The purpose of a robotic device is to mechanically assist the surgeon in carrying out the procedure according to the surgical plan and intentions of the clinician. Robotic advantages include high spatial accuracy and precision, non-fatigability, and tolerance of hazardous or difficult environments, including avoiding x-rays from fluoroscopy or maneuvering inside the bores of CT and MRI scanners. Furthermore, robotic dexterity allows a surgeon to have remote access or better access to constrained anatomic areas of interest, which are unreachable by standard tools. Thus, the goal of computer-integrated surgical systems is not to replace the physician with a machine, but rather to achieve consistent, precise, and improved medical outcomes by providing intelligence derived from *information* and controlled *action*, which augment the physician's natural capabilities.

## **2.2 Computer-Integrated Surgery Paradigms: Surgical Computer-Assisted Design (CAD)/Computer-Assisted Manufacturing (CAM) and Surgical Assistance**

Analogous to computer-integrated manufacturing, previous work<sup>58,59</sup> has classified computer-integrated surgical systems into two broad families: *surgical computer-assisted design (CAD)/computer-assisted model (CAM)* and *surgical assistants*. When referring to computer-integrated surgical systems, CAD depicts the process of planning and designing a model of the patient, while CAM denotes the registration, execution, and follow-up during the intervention. On the other hand, medical robotic systems are often synonymous with mechanical assistants that work cooperatively with surgeons. For example, a specific subclass of these *surgical assistants* includes tele-operated robotic systems. In contrast to direct control by the surgeon, tele-operated robotic systems

integrate computerized-control between input and output devices to augment or supplement the surgeon's ability. These categories, however, are not mutually exclusive. In fact, prior work<sup>60-62</sup> that is most similar to the described work are hybrid systems that integrate characteristics of both surgical CAD/CAM and surgical assistance.

### **2.2.1 Surgical CAD/CAM**

Early developments of the surgical CAD/CAM paradigm have been realized in neurosurgery and orthopedic surgery. The constant spatial anatomy of these two fields allows preoperative CAD models to serve as blueprints throughout the intervention, even after surgical modifications. Although both domains often involve operating in close proximity to deformable anatomy (e.g., the brain, spinal cord, and other neurovascular structures), their required geometric precision can be sufficiently represented in non-deformable models, constrained by rigid anatomy. For example, in neurosurgery, depending on the intervention, the brain's motion can be modeled as being constrained by the skull, although intra-cerebral movement is the subject of much research. Similarly, pedicle screw placement for orthopedic surgery can use guidance based on fixed bony anatomy (e.g., vertebrae), even though the spinal cord and musculature may deform during surgical intervention. Once registered, the accuracy of these systems relies upon the patient remaining rigidly fixed, as well as the geometric precision executed by the robotic arm using CAM, which allow surgeons to retain confidence regarding its execution.

Interaction between the clinician and the robot can range from minimal, with autonomous executions, to direct manual control. For example, the Programmable Universal Machine for Assembly (*PUMA*) 200 was used in 1985 to precisely manipulate

a biopsy cannula for neurosurgery<sup>63</sup>. Combined with a stereotactic frame in 1988, the *PUMA* became the first robot used on a human patient when it helped localize subcortical lesions during a brain biopsy<sup>64</sup>. Adaptations of the *PUMA* 560 for urology used a phantom model to support the transurethral resection of the prostate<sup>65</sup> and preceded the development of the *PROBOT* (Imperial College London, U.K.) in 1991.

In orthopedics, in 1986, collaboration between IBM's Thomas J. Watson Research Center and researchers at the University of California, Davis created an early prototype of the *ROBODOC* system, an innovative computer-guided system for total hip and total knee replacement surgeries. Research on *ROBODOC* began in the 1980s, when the conventional technique for hip and knee replacement surgery consisted of manual bone preparation guided by two-dimensional preoperative x-rays images. A surgical CAD created during preoperative planning was fabricated on a custom workstation that allowed the surgeon to position 3-D models of a commercial prosthesis with respect to the patient-specific CT volume. Intraoperatively, the surgical plan was transformed and registered to the robot coordinate system. The robot then machined the bone according to the plan, providing improved accuracy in both prostheses placement and bone removal. Preclinical *ROBODOC* experiments<sup>66</sup> demonstrated an order-of-magnitude improvement in precision over manual surgery. In 1992, the *ROBODOC*<sup>67,68</sup>, (previously Integrated Surgical Systems Sacramento, CA, U.S.A; now Curexo Technology Corporation, Fremont, CA U.S.A), became the first surgical robot to obtain an Investigational Device Exemption (IDE) from the Food and Drug Administration (FDA). As of 2014, the *ROBODOC* surgical system is the only active robotic system cleared by the FDA with

several studies showing better fit, fill, and alignment for successful joint replacement procedures<sup>69</sup> being conducted around the world.

Subsequent introduction of several other robotic systems for joint replacement surgery include the Rio surgical robot<sup>70,71</sup> (previously Mako Surgical, Ft. Lauderdale, FL; currently Stryker, Kalamazoo, MI) and the Acrobot<sup>55</sup> system (previously Acrobot Company, Ltd, a spin-off from Imperial College London, U.K.; currently Stryker, Kalamazoo, MI). Similarly, several groups have recently proposed small orthopedic robotic attachments or completely freehand systems such as the NavioPFS<sup>72</sup> surgical system (Blue Belt Technologies, Pittsburgh, PA), which combines the control of a surgical burr with intraoperative navigation. The *NeuroMate* (previously Innovative Medical Machines International, Lyon, France; currently Renishaw, Inc., Gloucestershire, United Kingdom) is a robotic arm used in neurosurgery, often passively. It was designed to precisely hold tools at predetermined configurations in order to support delicate and accurate localization in reference to stereotactic frames. Coupled with an image-guided system, the *NeuroMate*<sup>57</sup> can accurately track needle placement within the surgical space. A comprehensive review of other surgical CAD/CAM systems is provided by Kazanzides *et al.*<sup>58</sup> and Taylor *et al.*<sup>59</sup>.

For surgical CAM, execution and verification of the preoperative plan is key. With rigid bony anatomy, it has been relatively easy to obtain image-based verification through CT and X-ray fluoroscopy. Other surgical CAM systems often employ optical and electromagnetic (EM) solutions, which track calibrated tools with respect to preoperative volumetric data. Application of these tracking systems can be found in arthroplasty robotic systems (RIO<sup>70</sup>, MAKO Surgical Corp, Fort Lauderdale, Florida), as well as

otolaryngology, and head and neck surgical<sup>73</sup> procedures. For example, in endonasal endoscopic skull base procedures<sup>73</sup>, neurosurgeons often use a StealthStation (Medtronic, Minneapolis, Minnesota) to ensure completion of delicate resection of pituitary lesions. Similarly, optical tracking systems have supported numerous clinical neurosurgical cases<sup>74</sup> with navigation provided by a BrainLab VectorVision system (BrainLab AG, Feldkirchen, Germany).

### **2.2.2 Surgical Assistance**

Dexterous surgical assistance robotic devices were designed to extend the limits of human mechanical capabilities: to improve accuracy, filter tremor, and reach remote or deep structures via a minimally-invasive approach. Enhanced instrumentation is especially desirable in minimally-invasive laparoscopic procedures, where the patient benefits from reduced trauma, especially when compared to the large incisions necessary for traditional open access surgery. In laparoscopy, long instruments are inserted through small incisions in the abdomen, creating a fulcrum effect, which inverts the motions of the surgeon and makes coordination difficult. Surgeons must adapt to a more complex workspace with limited 2D visualization and impairment of dexterity, even as the sense of touch is also diminished. In this environment, robotic devices can offer advantages, including high spatial accuracy and precision, non-fatigability, and tolerance of hazardous or difficult environments such as interventional imaging workspaces inside CT and MR scanner bores.

Since the 1980s, there has been a growing interest in the research and development of surgical robots to assist clinicians. In Europe, a collaboration of the Karlsruhe Nuclear Research Center in Karlsruhe, Germany, and the University of Tuebingen in Tuebingen,

Germany, produced the Advanced Robot and Telemanipulator System for Minimally Invasive Surgery (ARTEMIS<sup>75</sup> 1987), which used remote telemanipulators through “over the shoulder” hand input devices to provide manipulation capabilities.

The Society of American Gastrointestinal and Endoscopic Surgeons (SAGES) and the Minimally Invasive Robotic Association (MIRA) define robotic surgery “*as a surgical procedure or technology that adds a computer technology enhanced device to the interaction between a surgeon and a patient during a surgical operation and assumes some degree of control heretofore completely reserved for the surgeon*”. Computer technology integrated between the surgeon and patient can extend the geographical reach of healthcare, as well. In the early 1970s, the National Aeronautics and Space Administration’s (NASA) interest in providing surgical interventions for astronauts on missions spawned research efforts<sup>76</sup> in tele-surgical and tele-presence systems. The United States military’s interest in robotic surgery has sponsored the Telemedicine & Advanced Technology Research Center to cultivate research for deploying interventional systems in remote battlefields (e.g., the Trauma Pod<sup>77,78</sup>). Stanford Research Institute (SRI) developed the precursor to the *da Vinci* (see 2.2.2.1 *da Vinci* robotic system for more details) with a dexterous telemanipulator to greatly enhance vascular and nerve anastomoses for hand surgery<sup>79</sup>. At IBM in 1993, the LARS robot was developed as a remote-center-of-motion surgical assistant to hold an instrument or camera with a variety of human-machine interfaces and a smart controller<sup>80,81</sup>. The controller provided advanced image processing and display functions, as well as robot-control. Around the same time period, Computer Motion designed the Automated Endoscopic System for Optimal Positioning (*AESOP*)<sup>82</sup> and the *ZEUS* robotic surgical system. Preliminary

operations conducted with the *ZEUS* were in gynecology in 1998, followed by a beating heart coronary artery bypass graft in 1999<sup>83</sup>. Efforts in tele-medicine culminated in the Lindbergh Operation<sup>84</sup>, a trans-Atlantic cholecystectomy, performed in 2001. Tele-mentoring uses similar technology to create a virtual classroom, permitting a surgeon to remain at his/her hospital while instructing or proctoring a novice at a remote location. Tele-presence<sup>85</sup> thus provides a new strategy for the training of surgical residents.

#### **2.2.2.1 *da Vinci* Robotic System**

One of the more successful examples of surgical assistance in computer-integrated surgery can be found with the *da Vinci* Surgical System. Following FDA approval in 2000, the *da Vinci* has become a mainstream option in urological, gynecological, cardiothoracic, and numerous general surgical procedures<sup>86</sup>. Designed for tele-surgery, the *da Vinci* consists of a patient-side slave robot and a master control console. In a master-slave system, a human operator manipulates the master interface to generate movement of the slave device. The master interface tracks this action (e.g., transformation and forces) as it is passed to a control and communication layer, which are used to exchange information, thereby enabling tele-operation. The slave device follows the interpreted commands from the communication system and interacts with the remote environment from which relevant information on slave components are sent back to the controller.

The slave robot supports a stereoscopic endoscope and two to three dexterous surgical instruments (e.g., needle drivers, cautery scissors, forceps, etc.) held by cable-actuated arms. Each robotic arm is designed with remote center of motion (RCM) kinematics, resulting in an inherent safety regarding the spatial stability of the entry port.



As is the case in laparoscopic surgery, when compared to open surgery, the incisions for a *da Vinci*-assisted procedure are smaller, the risk of infection is less, and thus convalescence and hospital stays are significantly reduced. In fact, many robot-assisted laparoscopic studies have shown that the decreased pain, better cosmesis, and improved postoperative immune function result in decreased hospital stays and a quicker return to the workforce<sup>87,88</sup>. Retrospective cohort studies have shown that patients who have undergone robot-assisted laparoscopic hysterectomy have significantly lower chances of readmission, as compared with those who undergo laparoscopic, abdominal (open), and vaginal hysterectomy<sup>89</sup>. Similar benefits have been found in studies of robotic-assisted laparoscopic prostatectomy<sup>90</sup> and partial nephrectomy<sup>91</sup> for renal masses.

When compared with traditional laparoscopy, distinct advantages of the *da Vinci* Surgical System include a functional wrist at the end effector that supports the tools' full six DOF (degree of freedom), as compared to laparoscopic instruments with four DOF. The surgeon sits at the surgeon side console (SSC), where hand motions are transmitted through dexterous master manipulators. These motions are replicated by the slave manipulators and viewed through high-quality stereo displays of the SSC. High-fidelity 3D visualization of the surgical workspace is another significant improvement from standard 2D laparoscopic displays. The robotic system decreases complex hand-eye coordination through the refinement of tool motion and the inherent support of a remote center of motion for all instruments and camera, which are under the direct control of the primary surgeon. The process may also be more ergonomically and physiologically beneficial to the surgeon, because most of the procedures can be conducted in a comfortable sitting position and outside of the sterile field, which can be shielded from

intraoperative radiation. In this tele-robotic approach, the *da Vinci* relies upon the surgeon as an active sentry, ready to react to any unexpected event.

#### **2.2.2.2 Mixed Paradigm Systems in Minimally-Invasive Surgery**

With the advancement of robotic dexterity, miniaturization, and image magnification, new techniques in micro-surgical and macro-surgical interventions are now feasible in practice<sup>92,93</sup>. Motivated by better quality of life outcomes, clinical practice is rapidly replacing traditional open procedures with minimally invasive techniques. Nonetheless, minimally invasive surgery presents new difficulties for clinicians by greatly reducing their sensory capabilities. The first main challenge is the loss of the direct sense of touch, replaced by little to no feedback depending on tools and systems. Touch, or haptic feedback, allows surgeons to differentiate between soft and hard tissue, as well as pressure from fluids or pulse, which is essential for anastomosis. Secondly, these video-based procedures generate a transition from direct visual feedback to indirect, image-based feedback. Furthermore, reduced visual perception stems from a loss of depth in the case of monocular cameras, as well as a limited field of view (usually 70 degrees instead of 160 degrees for the human eye<sup>94</sup>).

Thus, enhancement of surgical skill is even more desired in light of diminished feedback for minimally-invasive approaches. Intelligence can be provided by combining mechanically-assisted *action* and medical *information* to create hybrid systems that exhibit *mixed paradigms* of surgical CAD/CAM and surgical assistance. Mixed paradigm systems have been applied in urology, where ultrasound-guided laparoscopic radical prostatectomy overlays virtual surgical margins<sup>95</sup>. Similarly, video augmentation in guided robotic surgery has been used in endoscopic skull base studies<sup>96-98</sup> and

laparoscopic partial nephrectomy<sup>99</sup>. For *da Vinci* interventions, many researchers<sup>100,101</sup> have used the *TilePro* input in the surgeon side console to inject ultrasound images displayed below the primary 3D view. Han *et al.*<sup>101</sup> further integrated intraoperative 3D ultrasound using 2D-3D registration. Fusion with ultrasound to CT can be found through overlays as mixed paradigm minimally-invasive systems used to guide laparoscopic needle ablation of a renal tumor<sup>95</sup>.

Mixed paradigm systems have also been applied towards biliary surgery<sup>47</sup>, where tracked preoperative CT provides input axial images that can be viewed in *TilePro*. Such studies show the usefulness of 3D depth perception in multi-interventional applications, including gallstone, colectomy, and sigmoidectomy. The role of depth was further emphasized when Herrell *et al.*<sup>102</sup> resected embedded targets from gel phantoms using the *da Vinci* robotic system. Using registered CT augmented with the location of a tracked tool tip achieved a resection ratio closer to the ideal, compared to the resection ratio from procedures without image guidance. Additionally, procedures using image guidance were shorter, averaging 8 vs. 13 minutes. Demonstration of the benefits of surgical CAD/CAM for *da Vinci*-assisted tissue resection supports the potential for clinical outcome improvements, such as the decreased removal of benign tissue while maintaining an appropriate surgical margin.

For visualization in MIS, high-resolution endoscopy/laparoscopy is a natural vehicle to display information and quantitative measurement in diagnostic and interventional medicine. In fact, fusion of medical information in conventional cameras naturally lends to *augmented reality*<sup>103</sup>. Indeed, various approaches to augmented reality have been explored to alleviate challenges of minimally-invasive surgery by

complementing the clinician's visual field with necessary medical information that facilitates task performance. Clinical applications of these complementary mixed paradigm systems in MIS are often referred to as augmented reality for image-guided robotic surgery (Section 5).

### **2.2.3 Chapter Summary and Discussion**

In minimally-invasive surgery, a physician is limited to video-based visualization and reduced tactile feedback, which create an environment where complex hand-eye coordination contributes to a high cognitive load. Preoperative plans, if brought into the operative suite, are traditionally viewed on external monitors separately from the intraoperative video source. Thus, a significant visualization gap exists between surgical plans and their execution. The surgeon must mentally register the information rendered in the triplanar views from preoperative CT/MRI with the video scene; this is especially challenging for inexperienced residents. In the case of *da Vinci* procedures, external images can be displayed within the surgeon's console through an interface called *TilePro*. However, their placement below the endoscopic video display requires constant diversion from the primary visual scene. In the context of increasingly surgical complexity in minimally-invasive surgery<sup>30</sup>, enhancements of the endoscopic video (i.e., the current, primary visual source) with enriched medical information may provide a more natural "window". Additionally, any initial registration between the *da Vinci* and patient image data requires updates once the surgeon begins manipulating tissue. Knowledge of the robotic kinematics can be applied to update tracking and registration, but currently this information is available only through an API that is not available on standard clinical systems. Therefore, in order to determine endoscope and tool positions, many prior

experiments have needed to incorporate additional tracking systems. Generally, minimally-invasive robotic systems can be sizeable and cumbersome systems that occupy precious operating room space while offering constrained workspaces. Furthermore, medical robots are very expensive and their training curriculum continues to evolve.

Historically, in order to address these challenges, researchers have developed computer-integrated systems capable of surgical assistance and computer-assisted design and modeling. In this dissertation, we proposed an intelligent, versatile image-guided robotic surgical system that enhances the physician's ability to better understand and navigate the spatial structure of the patient anatomy by taking advantage of both paradigms. This can be accomplished by augmenting the surgeon's visual field with patient-specific models derived from multi-modal medical images, while leveraging the mechanical dexterity of a modern robotic platform.

While many aforementioned groups have contributed various systems for image-guided robotic surgery, the work described here offers the following key distinctions:

For augmented reality in medicine, solutions offered by select groups<sup>61,104</sup> have required additional hardware, while other image-guided robotic systems<sup>100-102,105-107</sup> displayed enhanced visualization adjacent to the endoscopic display (e.g., in the *da Vinci TilePro* input). Additional equipment typically includes hardware for visualization and tracking devices. This is expensive, potentially cumbersome, and can impede the OR workflow and procedure time. In fact, Linte *et al.*<sup>108</sup> have designated the footprint of new technology as one of the major barriers to introducing augmented reality technology into the clinical environment.

In comparison, we propose video augmentation of the primary display of the

existing robotic surgical system (i.e., the stereoscopic viewport within the surgeon-side console of a *da Vinci* system), in order to directly bridge the gap between preoperative plans and operative visual scene. Furthermore, no additional hardware for tracking camera/tools transformation is added; instead, the described system uses kinematics from the *da Vinci* API, computer-vision, and C-arm based updates. Thus, a distinguishing factor of this effort is the avoidance of not only visualization hardware, but also the external intraoperative optical or electromagnetic tracking device used in prior studies<sup>61,109,110</sup>.

Furthermore, previous systems that have augmented *da Vinci* endoscopy<sup>62,109,110</sup>, have relied on preoperative data. In contrast, by integrating an intraoperative C-arm, this work addresses perioperative patient anatomy and setup. C-arm CBCT<sup>103,111-117</sup> have provided intraoperative imaging for guidance in surgery; however, intraoperative updates based on C-arm fluoroscopy and CBCT to address *da Vinci* surgical motion is unique to the work described here.

## 3 A Modular System for Image-Guided Robotic Surgery

In Section 1.3, we presented four distinct clinical applications to motivate the image guidance system presented in this dissertation. For three *robotic-endoscopic/thoracoscopic* interventions, namely Oropharyngeal Cancer, Cochlear Implants, Thoracic Wedge Resection, we demonstrate the modularity of our design by applying the same architectural instance of our proposed solution; however, for *endoscopic* endonasal treatment of skull base lesions, the architecture differs. The main differences, discussed in latter sections, can be found mainly in the visualization and navigation modules. However, the architecture’s ability to adapt and accommodate a variety of software and hardware interfaces, as required for different clinical scenarios, shows evidence of its versatility. As our exemplary clinical application, transoral robotic surgery (TORS) is more extensively emphasized in upcoming sections which will detail the architectural design, requirements and individual modules of the proposed system.

### 3.1 Clinical Workflow

An outline of the proposed clinical workflow (Figure 3.1) for our image-guided robotics system is presented below. Its adaptation and execution in preclinical

experiments for multiple surgical scenarios are described in Chapter 6. The main workflow steps are as follows:

### (1) Preoperative Planning

Commensurate with standard of practice, our solution creates a preoperative surgical plan based on diagnostic volumetric image data, such as CT and/or MR. However, as an extension of clinical practice for use in image guidance, we segment critical anatomical structures to generate models for augmented reality.

### (2) Intraoperative CBCT\*

In the operating room after the patient is prepared and situated for surgery, for select clinical applications such as TORS and thoracic robotic intervention, preoperative image data may no longer reflect accumulated perioperative deformations. Nonconformities from preoperative to perioperative setup for TORS include mouth retraction and oral tongue extension, whereas robotic thoracic interventions require lung deflation. For point-based registration, we first place radiolucent fiducials (~5-10 teflon spheres) on the surface of the target volume of interest (e.g., the oral tongue for TORS) before acquiring a CBCT image, prior to positioning the robot for intervention.

### (3) Registration

---

\* required only for interventions with significant changes between intraoperative position and preoperative acquisitions



Similar to related work<sup>118</sup> using ultrasound instead of CBCT, the registration between our preoperative surgical CAD/CAM and the patient requires a two-step process. First, a non-rigid spatial relationship is established between preoperative and intraoperative data. Second, a rigid registration, or affine transformation, aligns intraoperative data and video.

(a) Deformable Registration\*

From the CBCT data acquired in step (2), we segment anatomical landmarks (i.e. for TORS: hyoid, oral tongue; for thoracic: hemiazygos vein, pulmonary artery). The labeled oral tongue is used in a deformable CT to CBCT registration for TORS, detailed in Section 3.6.1.

(b) CT to Video Registration

From the CBCT acquired in step (2), we manually segment the registration surface fiducial markers. Now labeled in the imaging coordinate space, these point fiducials are also manually identified in robotic video coordinate space in order to establish the affine registration (Section 3.6.2), which aligns the surgical plan to robotic video.

(4) Augmented Reality

Video from standard endoscopy is captured and augmented in real time with overlays of critical structures segmented from preoperative data and registered through step (3).

(5) Updates

After initial registration, overlaid critical data from surgical CAD/CAM require updates that reflect camera motion. We track camera transformations by following robotic kinematics. For tumor resection and enhanced depth perception in TORS, our system updates tissue motion from resection and tool tracking using vision-based techniques. Additionally, for thoracic interventions, we acquire dual intraoperative 2D X-ray fluoroscopic images and perform 2D-3D registration for image-based updates from intraoperative 3D localization.

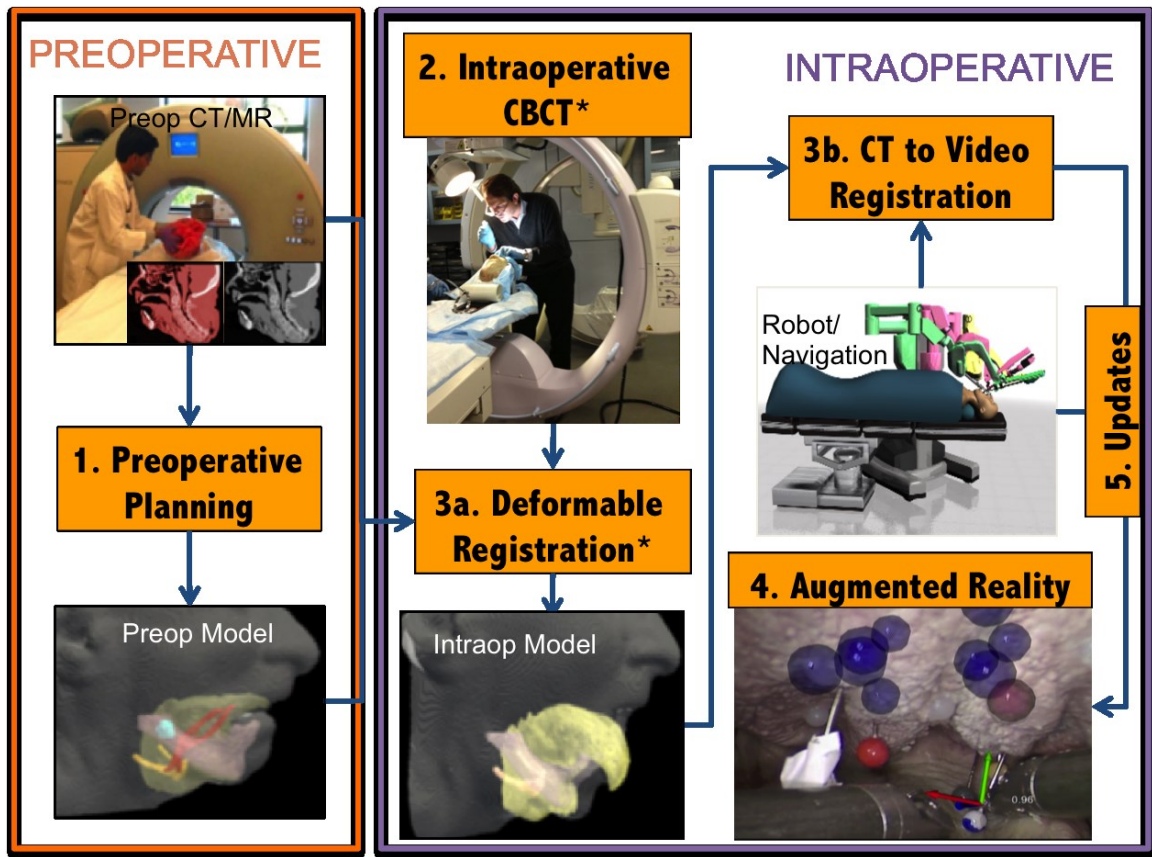


Figure 3.1 Image guidance workflow of a mixed paradigm system, featuring TORS as the main clinical motivation that requires deformable registration of surgical planning from preoperative CT to robotic stereoscopic video using intraoperative CBCT.

## **3.2 System Requirements**

### **3.2.1 Functional Requirements**

For intricate surgical tasks, inherent challenges in minimally-invasive surgery from a constrained workspace remain difficult, even with a robotic approach. Taught with an outside-in approach through open surgery, clinicians are accustomed to a natural field of view and using palpation to differentiate tissue gradations. However, with limited visualization, in addition to reduced force feedback, difficulties arise in determining anatomical landmarks and their boundaries. Under select surgical scenarios, particularly with the manipulation of delicate tissues and suture materials, the lack of tactile and haptic feedback is a significant disadvantage. Haptic feedback, even in the form of sensory substitution, has proven to facilitate the performance of skilled tasks, such as surgical knot tying<sup>119</sup>. To address these challenges, this work presents an image-guided robotic surgical (IGRS) system that combines information from preoperative planning, visualization from modern intraoperative imaging, and dexterity from robotic technical advancements.

For the mixed paradigm IGRS, as described in this document, the following modules and functionalities are required:

- Visualization Module

Using the visualization module, models from surgical plans are directly overlaid onto real-time capture of endoscopic images to create augmented reality for image guidance. This component therefore requires hardware that captures frames from monocular/stereoscopic live video streams. Imported

frames are fused with projections of the virtual scene, managed by the image guidance software, in order to create monocular/stereoscopic augmented video streams for output.

- User Interface Module

An interface for a technical operator, as well as the clinician, is used to select custom settings and conduct manual initialization tasks, such as data collection for registration and calibration. Our user interface options can be customized for user preference and include color of overlays, opacity, selection of critical data, and optional enhanced depth perception from novel views.

- Robotic/Navigation Module

Overlays for augmented reality require camera tracking in order to update the virtual scene. In addition, from tracked tool positions guidance, we incorporate feedback regarding tool end effectors as virtual objects for enhanced depth perception.

- Imaging System

Standard-of-care preoperative volumetric data is processed to create a surgical plan for image guidance. In select procedures where perioperative patient positioning requires an updated scan, we propose acquiring an intraoperative C-arm CBCT after the patient is set up for robotic intervention. In our solution, X-ray fluoroscopy can provide further intraoperative updates for target and tool 3D localization.

- Calibration

The development of several calibration functionalities was required in order to support the proposed system. First, optical camera calibration is needed to derive intrinsic and extrinsic parameters of the stereo endoscopes, used for creating the scene for virtual reality. Secondly, C-arm calibration, consisting of obtaining similar optical parameters, is necessary for 2D-3D registration. Additional robotic arm and manipulator calibrations are also required.

- Registration

In non-rigid surgical workspaces, a deformable transformation aligns the preoperative surgical plan and a perioperative C-arm CBCT. A rigid transformation between the registered surgical plan onto the video stream aligns the augmented reality to the intraoperative visualization.

### **3.2.2 Performance Requirements**

#### **3.2.2.1 Accuracy**

The reliability of an IGRS system must be assessed in order to provide the surgeon with a sense of how well the system is working. Ideally, a system is both accurate (i.e., the mean measurement that the system provides is very close to a reference true value) and precise (i.e., there is low variance in differences when the system returns the same measurement). In point-based registration, Fitzpatrick *et al.*<sup>120</sup> have formulated three error measurements that have been widely used in designating accuracy for image-guided surgery: (a) fiducial localization error (FLE), (b) fiducial registration error (FRE), and (c) target registration error (TRE). Point-based registration involves two sets of points,

moving and fixed; the derivation of the transformation from the former to the latter is formulated as the registration. FLE is defined as the distance between a localized point (e.g., segmented anatomical landmark from CT), and the point's ground truth location, which unfortunately is never known. FRE is the root mean square (RMS) error of the fiducial points used to compute the registration. In other words, after applying the registration transformation to the moving fiducial points, the RMS of the residuals of the registration is the distance between the transformed moving and fixed fiducials. We adopt TRE for evaluating the accuracy of different steps in our workflow.

Though accuracy requirements vary for the different clinical interventions motivating this body of work, we have focused on requirements for transoral robotic surgery (TORS) as an exemplary application. In base-of-tongue oropharyngeal cancer resection for TORS, the ultimate goal is to safely excise the tumor encased within a good margin of normal (negative) tissue. A good margin has been defined in a TORS context<sup>121</sup> as negative oncologic tissue ranging from 2 to 5 mm in radial thickness surrounding the tumor. Currently, the most utilized navigation systems for intraoperative image guidance consist of optical and electromagnetic (EM) solutions, which track calibrated tools to locate fiducials or sample surface points on the patient to use for registration. Previous assessments<sup>122,123</sup> of these systems in the clinical arena report accuracies ~2 mm target registration error (TRE).

TRE measures the RMS of distances of select targets in the transformed moving set of points to their corresponding fixed ground truth. In order to be comparable to conventional optical and EM trackers, we target our system to maintain an accuracy of 2 mm for each of the following registration steps: 3.6.1 Deformable CT to CBCT

Registration and 3.6.2 Rigid Video to CBCT Registration. First, these requirements are experimentally verified in Section 6.1.1 by applying deformable registration of CT to CBCT, using embedded targets on biofidelic phantom models. Second, a TRE for fiducial-based Video to CBCT registration is also measured at 2 mm (Section 3.6.2). Similarly, performance requirements for intraoperative updates from 3D localization using X-ray fluoroscopy from two C-arm views were also expected to be  $\leq 2$  mm (Section 4.3.4).

### 3.2.2.2 Visualization

Visualization is a key component in minimally-invasive surgery due to the challenges of a short baseline in a magnified view. A monocular overlay was achieved for *endoscopic* skull base surgery, adhering to standard *da Vinci* robotic systems; however, for the remaining robotic interventions, we presented a high fidelity representation of critical structures in stereo video augmentation. For augmented visualization in image-guided surgery, performance must have a low latency and high frame rate in order to approach the visual acuity adequate for the intervention at hand. In fact, while surgeons notice and can adapt with slower tool manipulations at high video latency (up to 100 ms), tele-surgical procedures show a decreased task performance with delays over 250 ms.<sup>124</sup>

Sources that contribute to the delay include hardware used for video capture whose performance consists of the following: shutter speed, camera digitization, data transfer, and processing time on an NVIDIA Quadro SDI graphic card that transfers captured frames to the image guidance software. Using a direct pass through to the output of the NVIDIA card (i.e., without injection of the overlay), we calculate the native latency of

these sources at  $\leq 90$  ms, which was estimated using a digital stopwatch with millisecond precision. This delay encapsulates the video latency of the original *da Vinci* cameras measured at  $\leq 57$  mm<sup>125</sup>, incremented with 33 mm buffering latency required for each interlaced frame as a limitation of the capture card. Programmatic stopwatches measured the direct pass through program at  $\sim 12$ -15 frames per second (FPS).

Basic augmented reality in medicine consists of a graphical projection showing select critical anatomies from the perspective of the current endoscopic view. Assuming an accurate registration, interpreting spatial localization within the projective plane (i.e., using the camera view) is straightforward. However, depth information (i.e., distances orthogonal to the camera view) is more ambiguous as it is interpreted from intuitive cues and perception. Thus, in order to eliminate ambiguities in measuring depth, features of the system included additional depth cues, including chromatic feedback, as well as numeric values of tracked tool tips using kinematic information from the robotic API as optional enhancements. In addition, by integrating an offline volume renderer, our system presented a novel view of the 3D scene, orthogonal to the endoscope and inlaid within the primary displays as a picture-in-picture. This design is therefore subject to the limitations of the performance rates of external devices, interface bandwidth, and applications. Depending on which features were active, the added latency of our overlaid image guidance, as estimated using the same method above, ranged from  $\sim 10$  to 60 ms. This caused our frame rate to drop to  $\sim 9$ -12 FPS.

The overall performance of our visualization was  $\leq 150$  ms in video latency and our expert surgeon was able to compensate for the high latency and low frame rate, but future work needs to explore both hardware and software solutions in order to address



these visualization performance limitations. Currently, the renderer of the image guidance system is updated through a single thread. Significant video performance improvements are expected to be achieved with multi-threading and/or parallel processing on GPUs, as well upgrades to RAM and CPU.

### **3.3 Architecture: A Modular Design**

An individual software component is a module that encapsulates a set of related functions and required data, therefore shielding details of its implementation while exposing specific services that other components can utilize. Component-based software engineering emphasizes the separation of functionalities or services. A versatile modular system for image-guided robot surgery integrating data from preoperative and intraoperative imaging was achieved with a generic component-based architecture, illustrated in Figure 3.2. This allows our approach to build loosely coupled independent components into an image-guided surgical system, reconfigurable for multiple clinical applications. With regard to system-wide coordination, components communicate with each other via subscription-based interfaces consisting of two types: *provided* and *required*. When a component offers services to the rest of the system, it adopts a *provided* interface as a signature. In order for component A to interface and run while connected to component B, A's *required* interface must be satisfied by those *provided* by B. Another important attribute of components is that they are *substitutable*, so that a new or optional component can replace or update an existing version without disrupting the remainder of the system. Key modules and functionalities of our design introduced in section 3.2.1 are detailed in the sections below.

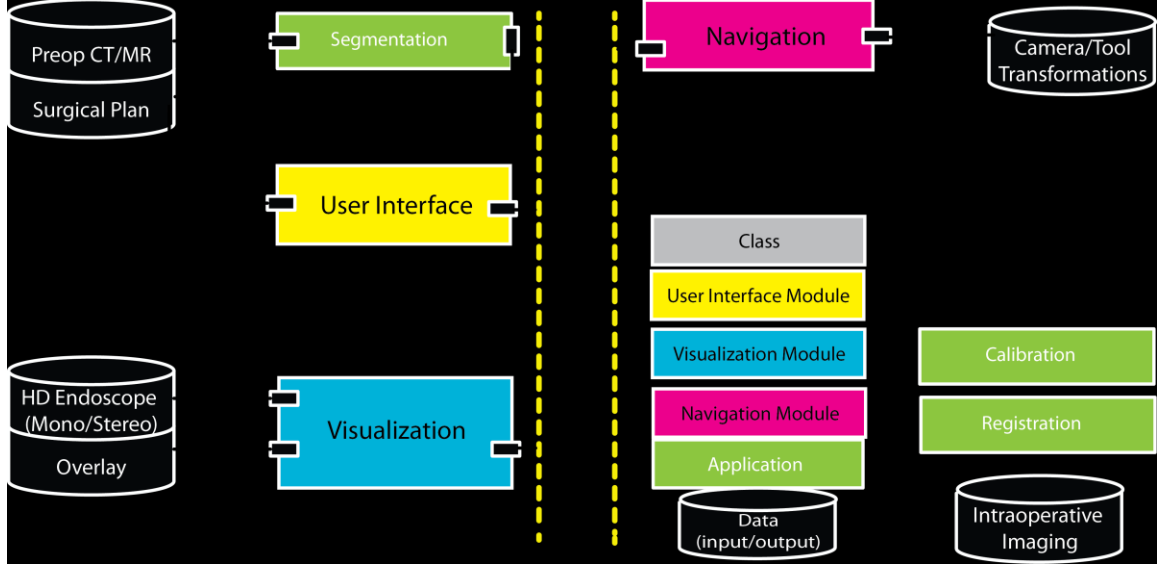


Figure 3.2 A diagram of the generic modular architecture of the proposed intraoperative image-guided robotic surgical system.

We adapt the generic system architecture (Figure 3.2) to two distinct image guidance systems: *robotic-endoscopic/thoracoscopic* image guidance system (Figure 3.3) and *endoscopic* skull base image guidance system (Figure 3.4) detailed below.

### ***Robotic-endoscopic/thoracoscopic Image Guidance System***

For the mixed paradigm robotic-endoscopic/thoracoscopic (i.e., *da Vinci*-assisted) image guidance system described in this dissertation, the following modules and functionalities are required:

- Visualization Module

Using the *cisst* stereo vision library (detailed in Section 3.3.1), dual video streams from *da Vinci* stereo endoscopes were captured for real-time augmentation using a research desktop computer. Fused with projections from preoperative planned data, augmented scenes replaced original stereo

endoscopic images directly on the surgical-side console. In addition, the augmented display also rendered a virtual orthogonal view (as a picture-in-picture), in order to provide enhanced depth perception.

- User Interface Module

The surgeon-side console (i.e., the conventional primary user interface for a *da Vinci* system) was extended with a graphical 3D user interface (detailed in Section 3.3.2). Functionalities included support for the surgeon to interact with objects in 3D virtual reality directly with the existing robotic manipulators.

- Robotic/Navigation Module

Camera transformations and tool positions derived from robotic kinematics were provided by the application programming interface (API) of the *da Vinci* system. The API was programmatically wrapped with A SAW *da Vinci* component that communicated information through the *cisst* multi-task layer to dependent components, such as the visualization module. For more details please refer to Section 3.3.3.

- Imaging System

For TORS, an intraoperative C-arm captured perioperative changes by acquiring a CBCT volumetric scan, after the patient was set up for robotic intervention. X-ray fluoroscopic images from the C-arm also provided intraoperative updates for target and tool 3D localization, explored for robot-assisted thoracic surgery.

- Calibration

Calibration steps included optical camera calibration and robotic arm calibration. In order to rely on kinematic information to track tool end effectors throughout the intervention for the primary robotic arm of the *da Vinci* patient-side cart (PSC), we computed a corrective transformation on the setup joints using forward kinematics. Additional calibration steps included modeling the X-ray imaging system as a pin-hole camera and determining C-arm parameters for rectification.

- Registration

First, in non-rigid surgical workspaces for TORS, a deformable transformation between surgical CAD/CAM and intraoperative C-arm CBCT was used to align preoperative planning to a perioperative setup. Second, a rigid point-based transformation was computed using segmented fiducials from intraoperative images in order to register the image guidance data to the video scene.

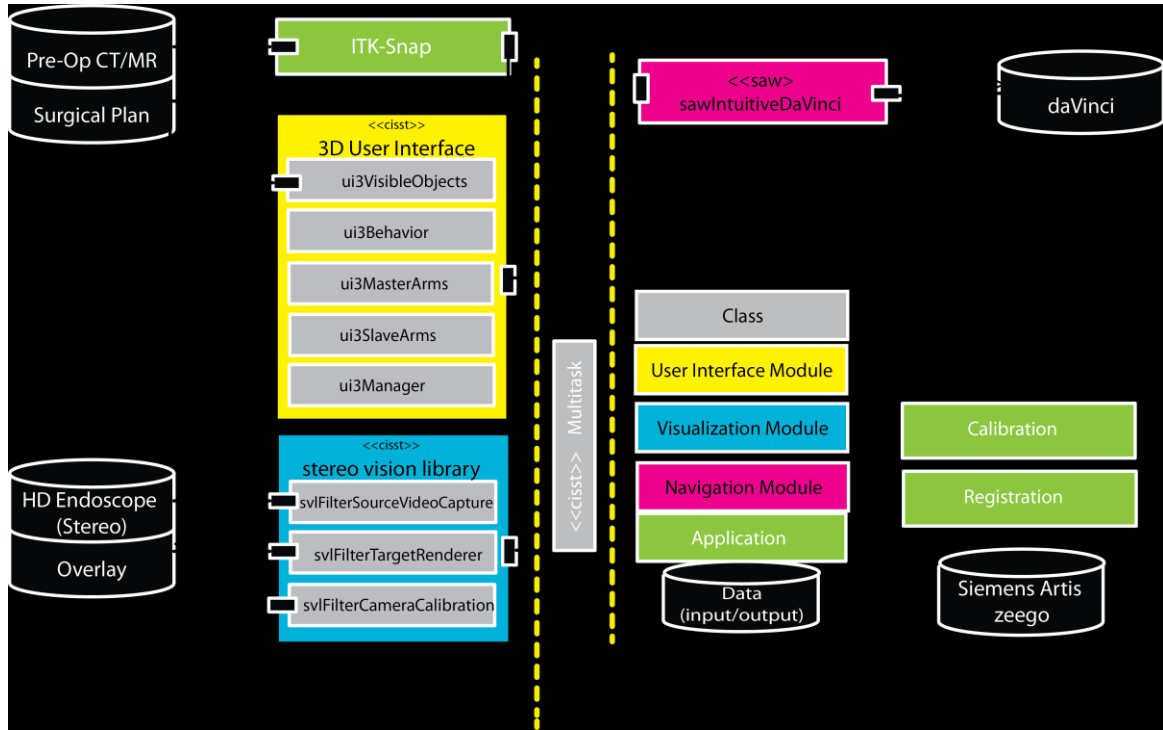


Figure 3.3 An architectural component diagram detailing the libraries supporting stereo video augmentation and modular design of the proposed intraoperative image-guided robotic system used for TORS.

### ***Endoscopic Image Guidance System***

For the endoscopic endonasal skull base image guidance system, as described in this dissertation, the following modules and functionalities are required:

- Visualization Module

The *cisst* stereo vision library was used to process monocular video from a HD Karl Storz (Karl Storz Inc., Tuttlingen Germany).. The image guidance system overlaid models of critical skull base structures (e.g., carotid arteries, optical nerves, pituitary lesions), derived from standard preoperative CT/MR images onto the captured frames.

- User Interface Module

TREK<sup>126</sup>, a custom image guidance software built from 3D Slicer, was modified to create a light-weight interface for image-guided skull base surgery.

- Robotic/Navigation Module

A commercial navigation system (StealthStation, Medtronic Inc.) was used to track camera transformations. The *cisst* c++ multi-task library was wrapped with python (SWIG) to handle inter-module communication.

- Calibration

To create the virtual scene, optical camera calibration derived intrinsic and extrinsic parameters of the monocular endoscope. Hand-eye calibration resolved the unknown transformation of the endoscope's tracked rigid body to the optical center.

- Registration

Standard preoperative CT to patient registration, captured by the StealthStation, was reused by this image-guidance system to align the virtual surgical plan to video.

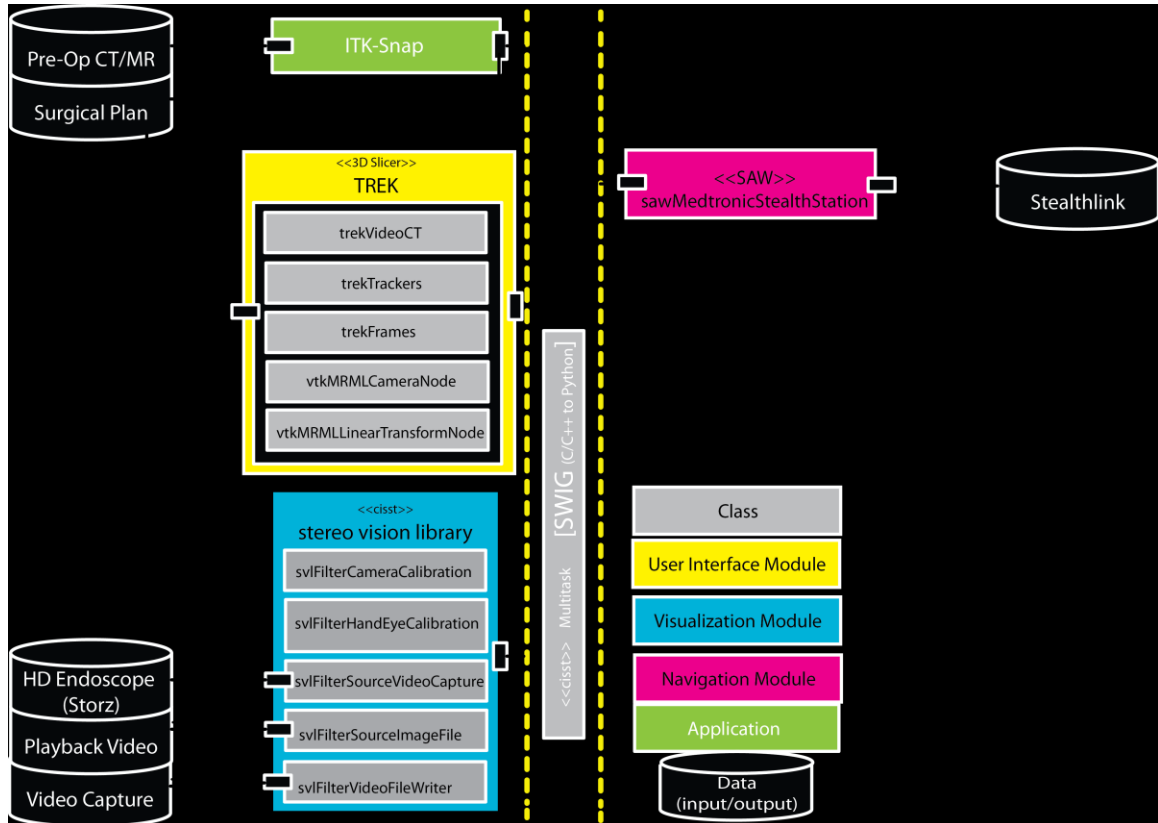


Figure 3.4 An architectural component diagram detailing the classes supporting video augmentation and emphasizing the modular design of the system architecture. The system is an extension of the TREK architecture for image-guided surgery, binding *cisst*/SAW libraries for real-time tracking and registration with 3D Slicer libraries for front-end visualization. The specific embodiment described in this dissertation was intended to streamline calibration processes in a manner suitable to clinical use by a trained OR technologist without disruption of OR workflow.

### 3.3.1 Visualization Module

For a mixed paradigm image-guided robotic surgical (IGRS) system, the visualization module is a critical component because it dominates the user experience. Unlike other related work in IGRS<sup>100-102,105-107</sup> that have integrated information from surgical plans for the *da Vinci* through TilePro or with additional displays<sup>76,127,128</sup> (e.g., head-mounted devices and Google glass), we directly augment the primary endoscopic video. When fusing virtual information, it is critical that any augmentations do not overwhelm the field of view and detract from the procedure. We address this by overlaying select critical anatomy, as opposed to dense volumetric rendering, in addition to customizing user preference for color and opacity as detailed below.

Video augmentation was achieved within the modular architecture, illustrated in Figure 3.2, by extending the *cisst* stereo vision library (SVL) from the *cisst/SURGICAL ASSISTANCE WORKSTATION (SAW)* open-source toolkit<sup>129</sup>, developed at the Engineering Research Center for Computer Integrated Surgical Systems and Technology (CISST ERC, Johns Hopkins University, Baltimore, MD). A research desktop was built using a Dell Precision T7500 with dual 6-core Intel Xeon processors, running dual boot Ubuntu Linux and Windows. Within the SVL library, a base class of video capture devices interfaced with various graphics/video capture cards compatible to each clinical application.

For *da Vinci* procedures, video frame grabbers processed dual HD-SDI SMPTE 274 (1080i@59.94Hz) video signals from the stereo camera control units, computed virtual scene updates, and injected augmented dual HD-SDI channels back into the visual core of the robotic system. Initially, two Matrox Vio (Quebec, Canada) PCIe video



capture cards with overlay input and passive video throughput served as the image processing hardware. However, these did not have native support for transparency, so we first had to stipple the overlaid mesh. The Vios were replaced with a single Nvidia Quadro SDI Capture Card (Santa Clara, CA, USA), allowing for a more natural blend of the stereoscopic video and image guidance through active dual stream alpha-blended overlays.

For skull base procedures, the camera control unit from the monocular high-definition Storz endoscope (1080i@59.94Hz) was accessed through a DVI-I port. A Gefen EXT-DVI-2-HDS DISSL system converted this signal to HD-SDI SMPTE 274. We extended the SVL library to support a subset of Blackmagic Decklink devices to capture the HD-SDI endoscopic video for the image guidance software.

### **3.3.2 User Interface Module**

Similar to visualization, for *da Vinci* interventions, we enhanced the surgeon-side console (SSC) directly, unlike other prior work where additional equipment<sup>47</sup> was required, to manipulate the 3D CAD/CAM data. Currently, unlike 2D interfaces, there is no best practice for three-dimensional user interfaces using the WIMP paradigm (windows, icons, menus, and pointing). Bowman *et al.*<sup>130</sup> give a comprehensive introduction to 3D user interfaces and detailed information on why 3D interaction is difficult, pointing out that classic 2D computer interaction paradigms such as windows, mouse pointer, menus, and keyboards do not generally translate well for 3D displays and environments.

Within the stereoscopic environment of the *da Vinci*, the surgeon is accustomed to 3D visualization. Prior work has demonstrated the advantage of 3D visualization and

interfaces in orthopedic application. For example, Traub *et al.*<sup>131</sup> showed that surgeons were able to perform drilling experiments faster with *in situ* 3D visualization compared to a navigation system with a classic 2D display. Thus, the user interface we created is rendered in stereo as a 3D virtual menu embedded within the augmented scene. Critical data (i.e., for TORS: the tumor, lingual/carotid artery, lingual nerve, etc.) were manually segmented from preoperative CT/CBCT images using ITK-Snap and saved in Visualization Toolkit (VTK)<sup>132</sup> formats. The VTK data structures are loaded into the *cisst* 3D user interface library (UIL), which supports additional objects for 3D rendering and *behaviors*, classes that process user inputs to update the augmented 3D scene. In a special user mode, termed *masters-as-mice*, the surgeon is able to access and manipulate the virtual CAD models through master manipulators (MTMs) on the SSC. In this mode, the clutch pedal decouples the MTM from the PSC and virtual 3D cursors (Figure 3.5, white spheres) follow the movement of the MTMs, allowing for dynamic user input. Using this interface, the physician can interact and manipulate the overlaid objects in 3D, as well as select menu options to change their color, opacity, and visibility.

The skull base image-guidance system extended TREK<sup>126</sup> (Figure 3.6), a software architecture for image-guided surgery, to create a clinically deployable interface. TREK binds open-source libraries for image visualization and analysis from 3D Slicer<sup>132</sup> (NAMIC kit, Brigham & Women's Hospital, Cambridge MA). Our efforts integrated real-time tracking and registration from a commercial navigation system, Medtronic StealthStation, detailed below.

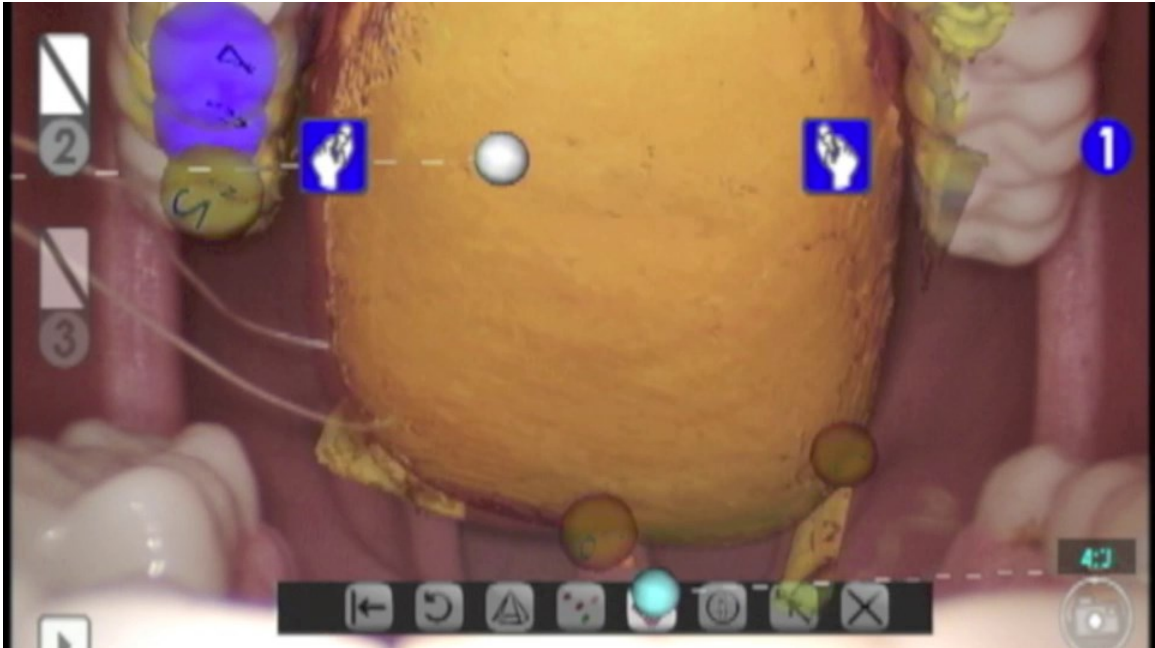


Figure 3.5 Visualization and 3D user interface for transoral robotic surgery (synthetic head phantom).

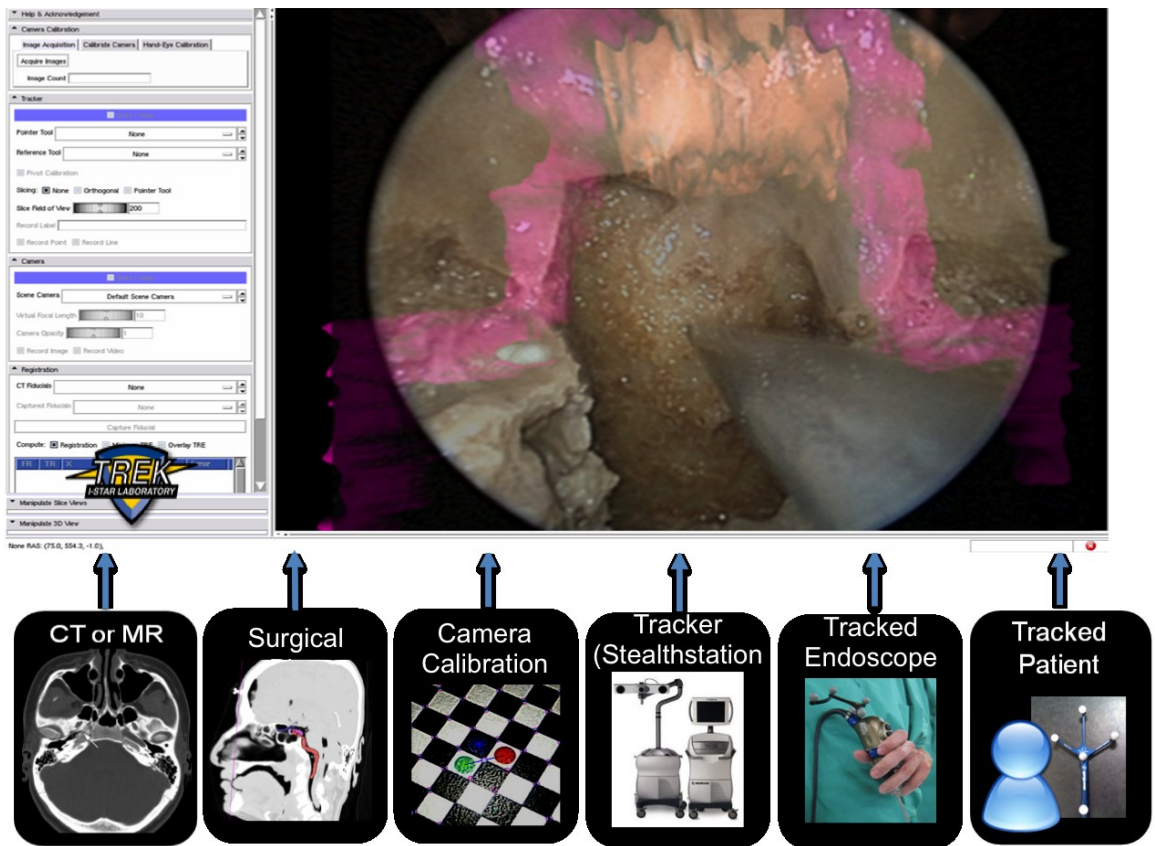


Figure 3.6 Visualization and user interface for endoscopic endonasal skull base surgery using TREK and Medtronic StealthStation.

### 3.3.3 Robotic/Navigation Module

Throughout an intervention, knowledge regarding camera movement and relative tool localization (explored for TORS) are used to update the augmented virtual scene. For our clinical applications, this information source has been either a Medtronic StealthStation (Medtronic Inc., Minneapolis MN) or a *da Vinci* system. By extending these current standard-of-care systems, we minimize clinical disruption and take advantage of user familiarity.

Through a collaboration between the Engineering Research Center for Computer-Integrated Surgical Systems and Technology (Department of Computer Science, Johns Hopkins University) and Intuitive Surgical, Inc., the IGRS system presented here was able to interface with the *da Vinci* platform via a research API. Connections between a UIL and SAW component, which wraps the API, *behaviors* are established and supervised by a component manager, a class of the *cisst* multi-task (MTS) library. In the MTS framework, which utilizes component-based software engineering concepts, all interaction occur via *provided* and *required* interfaces.

Therefore, the SAW *da Vinci* component *provides* robotic system information, such as clutching, camera movement, master manipulator movement, and resultant real-time joint positions of patient-side manipulators (PSMs), to callback functions *required* by *behaviors* of the 3DUI library. This architecture decouples the input/output device (*da Vinci* SSC and PSC, respectively) and the visualization software into separate components exposed through a multi-threaded communication layer. Such a modular

design supports modifications within elements and the exchange of entire components for flexibility and extendibility.

An overview of the main components and features of the *da Vinci* robotic system was given in Section 2.2.2.1. In the sections below we highlight key differences between models.

### **3.3.3.1 *da Vinci S***

Released in 2006, the *da Vinci S* updated its predecessor, the *da Vinci Classic*. Functional differences included increases in tool shaft length, tool flexion, and an increased inertia when the tools are clutched. Major structural changes were revealed with the robotic arms on the patient side cart, where a streamlined design improved the range in workspace of the robot and avoidance of collisions. As reported through the API, the range in motion of the robotic tools constrained the workspace requirement of our image guidance system. Although robotic kinematics and interior controller modifications affected the intrinsic systematic design, Kwartowitz *et al.*<sup>133</sup> conducted a series of phantom tests and determined that the mean localization errors of the *classic* compared to the *S* was relatively similar at 1.02 mm and 1.05 mm, respectively.

Changes directly impacting our proposed solution can be found with a new user interface that featured stereo high definition (HD) cameras and an external display for the patient side assistant. The *S* model supports both a standard and a high definition (HD) stereo 3D camera head. For the purposes of our efforts, we developed our visualization module and camera calibration functionalities to be useable with a stereo HD Ikegami camera control units with an 8.5 mm, zero degree endoscope as supported by the *da Vinci S*.

### **3.3.3.2 *da Vinci Si***

Following the *S*, Intuitive Surgical Inc. released the *da Vinci Si* in 2009. While the structural exterior of the patient-side console largely remained unchanged, differences between the two systems can be seen in upgrades to the surgeon-side console. Improved usability features included multiple ergonomic adjustments and an integrated touchpad that allowed the user to set video, audio, and ergonomic settings through a comprehensive digital menu. In addition, switches on the master manipulators supported finger-tip control for clutching and adjusting endoscopic focal length. As described thus far, these features were supported on the *S*; however, on the *Si*, they were now controlled with an updated interface, and therefore were accommodated by our system as new API events that correspond to existing functionality.

On the other hand, the *Si* release enabled a new configuration that allowed two surgeons to collaborate during a procedure for *da Vinci*-enabled surgical assistance, or facilitate teaching through two synchronized surgeon side consoles. Supporting features include exchanging control of the patient side manipulators and the endoscope, as well as intercom communication. To support these new features of the *Si*, our system was extended to support a SAW *da Vinci Si* component as an alternative to the *S* component. In addition, experiments with the *Si* included accommodating zero and 30 degree endoscopes both at 12 mm and 8.5 mm.

### **3.3.3.3 *da Vinci Sp***

The *da Vinci Sp* Surgical System, Model SP999 (Figure 3.7), is a research prototype that attained FDA 510(k) clearance in April 2014. As indicated through released reports from Intuitive Surgical, Inc., anticipated efforts include product

refinement and optimization fully compatible with the most recently released model, the *da Vinci Xi* Surgical System, which will require additional regulatory clearances.

This research model expands Intuitive Surgical, Inc.’s product capabilities in single incision robotic laparoscopic surgery as a significant departure from multi-port systems such as the *S* and *Si*. Using just one 25 mm cannula, this is a single incision *da Vinci* system with an articulating stereo HD camera and three fully articulating tools, housed by the Entry Guide Manipulator (EGM). In single-port *da Vinci-assisted* surgery, when compared to the *da Vinci* Single-Site approach, the wrists of the *Sp* instruments have two more degrees of freedom (DOF) than the passively flexible Single-Site instruments, which are not wristed. While the variety of EndoWrist *Sp* instruments continues to be developed and refined, current research surgical tools include: flexible endoscopes, blunt and sharp endoscopic dissectors, scissors, forceps/pick-ups, needle holders, endoscopic retractors, and accessories for endoscopic manipulation of tissue, including grasping, cutting, blunt and sharp dissection, approximation, ligation, electrocautery, and suturing.

Although the initial FDA clearance is specific to adult urologic surgical procedures that are appropriate for a single port approach, we took advantage of the streamlined design to validate the expandability of our architecture by exploring additional clinical applications in gynecology and general surgery (Section 4.3.1). Furthermore, in Section 5.2.2 we describe the necessary changes in the visualization module required for updates to video augmentation using the *Sp*. Modifications addressed novel articulated camera motions and synchronized camera and tool movements. Significant changes of our

system from the interfaces that supported  $S/S_i$  included forward kinematics of the flexible endoscope using Denavit-Hartenberg parameters.

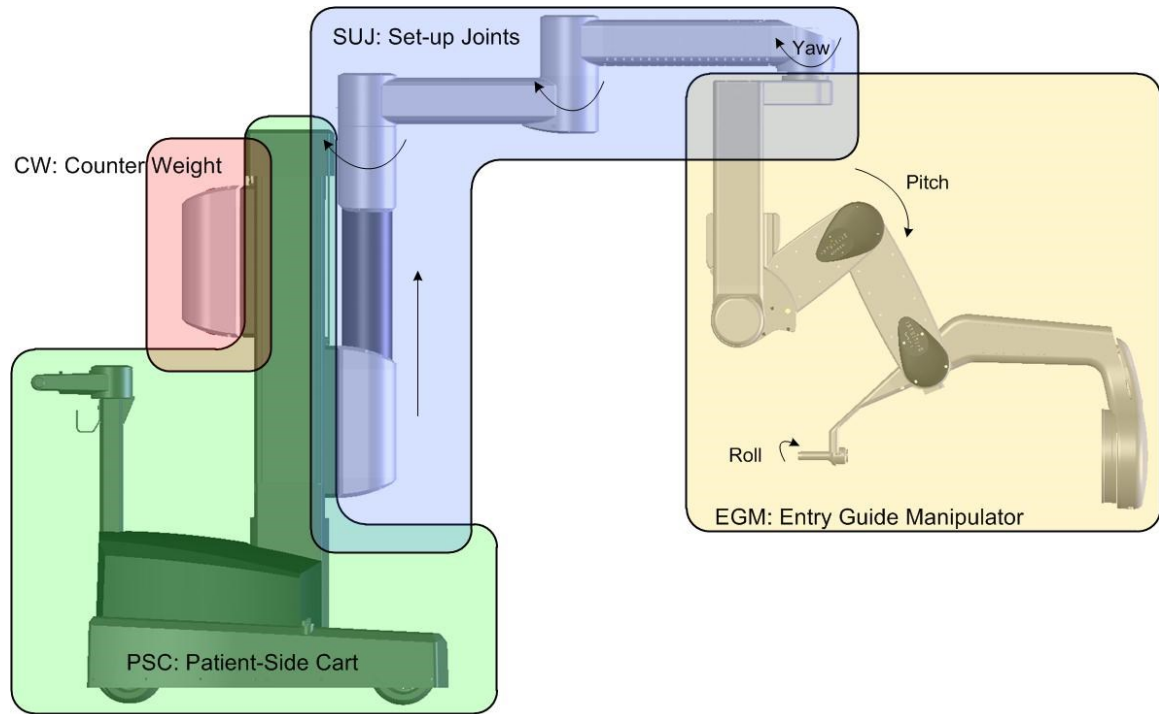


Figure 3.7 Labeled components of the research system, *da Vinci Sp*.

### 3.3.3.4 Medtronic Stealthstation

In skull base surgery, we further extended SAW by developing a new component for *Stealthlink*, the research API for the Medtronic StealthStation). This interface to the commercial Medtronic navigation system provided novel infrared tracking of a rigid body marker attached to the endoscope, tracking of proprietary pointers, and a reference marker attached to the stereotactic head frame. In addition, we were able to directly use the registration of preoperative CT to patient, a standard step in the conventional clinical workflow, to align surgical CAD/CAM from preoperative CT to video.



### **3.4 Imaging System: C-arm CBCT, Angiography, and Fluoroscopy**

Despite the wealth of information available due to the advances of technology in medical imaging, the extent of preoperative diagnostic data readily integrated during robotic surgery is still limited. Diagnostic images are used to examine the patient's anatomy, devise a suitable treatment, and plan for its execution. The most common imaging modalities used preoperatively for surgery include CT and MR imaging.

Though these high quality volumetric data sufficiently delineate critical anatomical boundaries for image guidance, they do not address perioperative setup and the current state of the patient. When setting up for robotic intervention, the patient is likely in a different physical arrangement than during the preoperative data acquisition. Common changes include: lateral rotation, insufflation, and retraction; also, for thoracic interventions, the lung is collapsed. In natural orifice approaches, such as TORS, neighboring tissue is retracted and displaced to create workspace and access. Furthermore, a cancer patient undergoing chemo-radiation therapy in addition to other health-related stress may exhibit dramatic physical changes between the operative day and diagnostic imaging. Therefore, intraoperative imaging, which captures these changes, can be invaluable for an updated high fidelity surgical CAD/CAM. Challenges for acceptance of new intraoperative technologies include performance requirements where modalities must operate in near real-time and be physically compatible with the standard operating room. Image acquisitions are expected to provide highly accurate guidance at sufficient spatial resolution with minimal impact on the normal interventional workflow. Currently, common real-time intra-operative imaging modalities include ultrasound (US)

imaging<sup>51,100,134,135</sup>, X-ray fluoroscopy<sup>52,136,137</sup>, and more recently, cone-beam CT<sup>103,126,138-140</sup>, and intra-operative MRI<sup>141-143</sup>.

Fluoroscopy is a real-time and cost-effective modality, though image quality can vary greatly depending on the age and technology of the system. C-arms in clinical use range from older systems that use image intensifiers to motor-actuated flat panel detectors that are synchronized with an x-ray source. Thus, C-arm image capabilities vary from distorted single 2D planar x-rays images to 3D reconstructed volumes (e.g., from CBCT). The main C-arm system that we have integrated for this IGRS system is the Siemens Artis *zeego*<sup>144</sup>, which consists of a flat panel detector (30x40 cm) synchronized with an x-ray tube mounted to a 6-DOF of freedom robot. However, for comprehensiveness and to test the flexibility of our proposed workflow, we also examined the possibility of using an OEC 9600 C-arm (1998, tri-mode 12/9/6" image intensifier), a representative older model in the current spectrum of C-arm technology. Unlike the *zeego*, the OEC only provides 2D radiographic images and is not capable of CBCT, the reconstruction of a volume from a series of x-ray projections. Thus, its integration was explored only for 2D-3D registration and 3D localization from dual projections (Section 4.3.3). Using the *zeego*, we explored intraoperative x-ray tomography, CBCT, and CBCT angiography (CBCTA). Furthermore, intraoperative volumetric C-arm images can be used to fuse additional preoperative modalities<sup>145</sup>.

For our main clinical motivation, TORS, a CBCT image acquired immediately prior to robotic docking captures perioperative deformation of the oral workspace (e.g., from neck flexure, tongue extension, mouth retraction). In thoracic intervention, a perioperative CBCT can capture gross changes from patient positioning from a lateral

rotation and lung collapse. Although deflated lung tissue will exhibit low contrast, intraoperative details regarding critical vasculature in real-time can contribute to precision in registration and visualization, which improves the safety of the patient. In Section 6.2, CBCTA is extensively examined in states of inflation and deflation for intraoperative thoracic application, as well as analysis of displacement from the perioperative setup. We explored the challenging cases of angiography in hepatic applications, where arterial and venous portal phases (Figure 3.8) are timed with contrast injection and reconstruction from CBCT to illuminate the major vessels interrogated during partial nephrectomies.

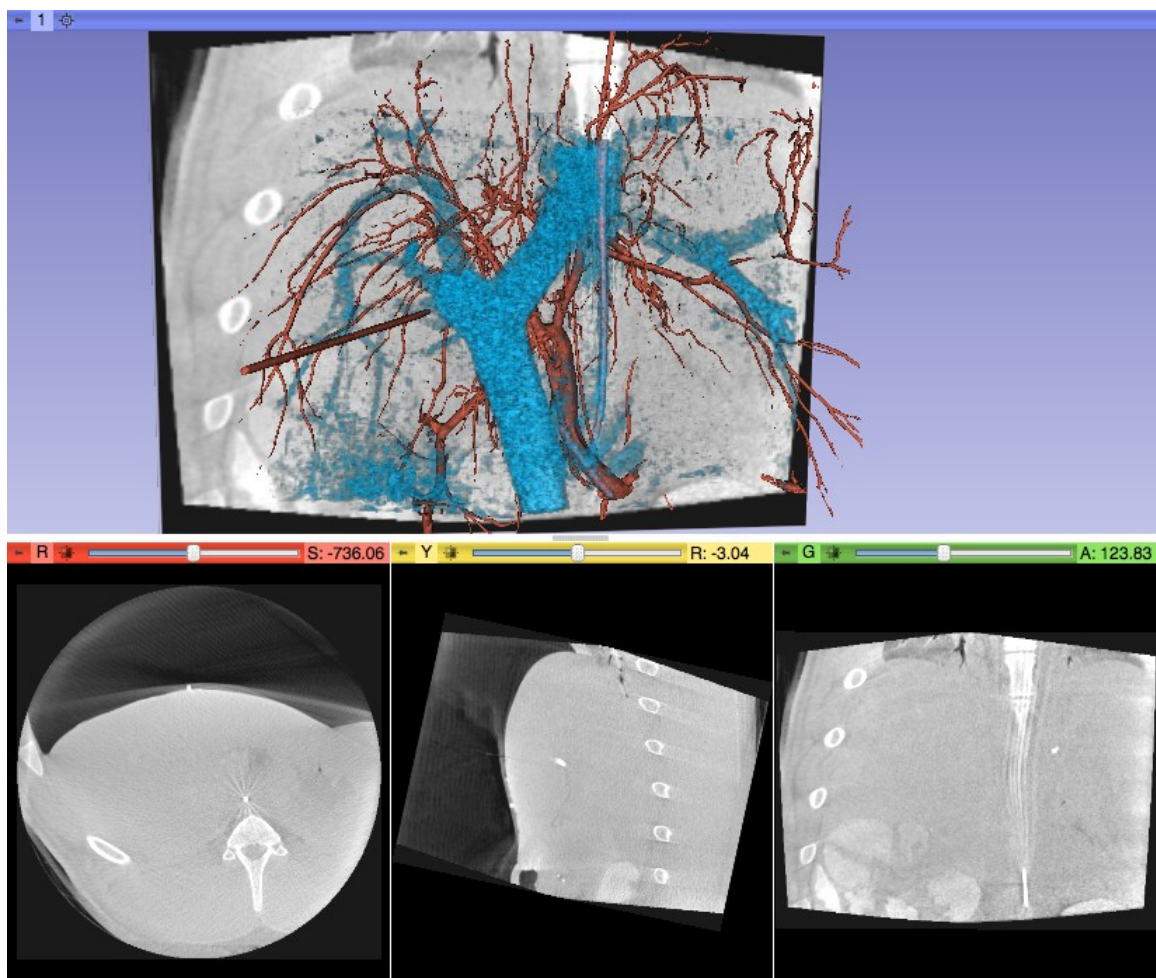


Figure 3.8 Artis zeego CBCT angiography of a porcine liver with a synthetic tumor (white in triplanar views), arterial (red), and venus portal (blue) phases.

### **3.5 Calibration**

Calibration is generally a process executed to determine characterizing parameters of a system or device. An acceptable range of values for these characterizations may be known, but calibrating can optimize and more exactly display the true values used by dependent applications. For example, standard camera calibration determines the intrinsic and extrinsic parameters of endoscopes, which is necessary for virtual cameras used in augmented reality. Furthermore, distortions from optics as well as magnetic fields for C-arm radiographs can be corrected using rectification parameters determined using the techniques explained in the sections below.

#### **3.5.1 Optical Camera Calibration**

Determining the optical parameters of the endoscope is essential for our proposed image guidance, which uses high fidelity overlay of augmented reality from medical images. This process is a ubiquitous step known as optical camera calibration. In camera calibration, we determine two sets of parameters, the intrinsic and extrinsic. Intrinsic parameters consist of the elements of the optical lens, including the focal length, optical center, and distortion characteristics. Extrinsic parameters result from a rigid transformation of the optical center of the camera with respect to a known global coordinate system.

Previous work in camera calibration has been available through the DLR toolbox<sup>146</sup> (DLR CalLab and CalDe, German Aerospace Center, Wessling, Germany).

However, using this freeware requires ~40 min for camera calibration for standard-definition video. The DLR process is lengthy due to manual identification of the origin and orientation for each image, comparable to that of the MATLAB (v2011b, The Mathworks, Natick MA) camera calibration toolbox (Camera Calibration Toolbox for Matlab, Caltech, Pasadena, California), in which the users repeatedly identify the same four corners of the calibration grid. Although faster (on the order of minutes for an experienced user), the MATLAB toolbox was found to be less accurate in computing barrel distortion correction where the calibration grid was not always entirely in the field of view.

To streamline this step for a fast deployable clinical workflow, we developed a semi-automated camera calibration process by extending functionalities provided in OpenCV<sup>147</sup> (v2.1, Intel Research/Willow Garage Inc., Menlo Park CA). An image filter class was implemented in SVL where eigenvalue-based features were correlated with iterative homography and used to solve for the intrinsic and extrinsic camera parameters. We eliminated the need for manual identification of the origin and orientation of a custom checkerboard by automatically segmenting embedded red, green, and blue markers to resolve the local coordinate system of each image.

### **3.5.2 Hand-Eye Calibration**

For navigation in endoscopic skull base surgery in select standard clinical cases, a Medtronic StealthStation can be used to provide registration and tracking with respect to preoperative CT scans. We enhanced this approach by deploying a system for 2D video augmentation, interfacing with the StealthStation to provide optical tracking of a rigid

marker attached to the endoscope. Registration of endoscopic video with preoperative 3D CT image data,  $^{Endoscope}T_{CT}$ , was derived by the following transformations:

$$^{Endoscope}T_{CT} = (^{Endoscope}T_{rigidBody})(^{rigidBody}T_{tracker}) (^{tracker}T_{reference})(^{reference}T_{RAS})(^{RAS}T_{patient})(^{Patient}T_{CT}) \quad (3.1)$$

Markers on a stereotactic frame or radio-opaque fiducials in CT provide the  $^{Patient}T_{CT}$  transformation. As part of the standard-of-care registration step, either the same fiducial markers or surface point sampling was used to localize points with a tracked pointer to determine transformation from the system coordinates in Right, Anterior, Superior (RAS) to a reference,  $^{reference}T_{RAS}$ . Additional transformation from RAS to patient as recorded by the image acquisition system resolves the transformation from StealthStation to the image guidance system, thus reusing the exact transformation obtained for the existing patient registration process.

The final transformation,  $^{Endoscope}T_{rigidBody}$ , is the unknown relationship of the tracked rigid body to the optical center of the endoscope. Solving for  $X$ , the homogeneous transformation from camera to endoscope during calibration by exploiting the relationship of multiple camera poses is a process known as hand-eye calibration. In a single camera pose,  $A_i$  gives the transformation matrix from the tracker to the rigid body, while  $B_i$  is the transformation matrix from the optical center to the calibration grid, derived from camera calibration. One motion (i.e., two poses) yields the conventional hand-eye equation:

$$AX = XB \quad (3.2)$$

where

$$A = A_1^{-1}A_2 \text{ and } B = B_1 B_2^{-1} .$$

The hand-eye calibration for the skull base image guidance system applies

compact dual quaternions, the algebraic counterpart of screws as proposed by Daniilidis<sup>148</sup>. A quaternion,  $(s, \vec{q})$ , is a 4-tuple representation of a rotation extending complex numbers to  $R^4$ . Dual quaternions,  $(\check{s}, \check{\vec{q}})$ , extend quaternion representation with  $\check{s}$  as a dual number, and  $\check{\vec{q}}$  as a dual vector. A line in space with direction  $\vec{l}$  through a point  $\vec{p}$  can be represented with the six-tuple  $(\vec{l}, \vec{m})$ , where the line moment,  $\vec{m}$ , is equal to  $\vec{p} \times \vec{l}$  and can be denoted with four parameters; together with the rotation angle  $\theta$  and the translation along the pitch  $d$ , these parameters constitute the six degrees of freedom of a rigid transformation. A rigid transformation can therefore be modeled as a rotation with the same angle about a line in space (i.e., the screw axis) that does not pass through the origin and a translation along this axis. The direction  $\vec{l}$  is parallel to the rotation axis, and the pitch  $d$  is the projection of the translation on the rotation axis. Using these relations, Daniilidis proved that the hand-eye transformation is independent of the angle and the pitch of the camera and hand motions, and depends only on the line parameters of the screw axes. With dual quaternions characterizing our transformations, we have:

$$\check{\alpha} = \check{q}\check{b}\check{q} \quad (3.3)$$

where  $\check{\alpha}$  denotes the extrinsic camera parameters,  $\check{b}$ , the tracked endoscope, and  $\check{q}$ , the  $Endoscope T_{rigidBody}$  transformation. Hand-eye calibration with dual quaternions provides a fast and efficient simultaneous solution of rotation and translation using singular value decomposition.

### 3.5.3 C-arm Calibration

X-ray images acquired using image intensifiers exhibit several characteristic distortions due to both external and internal factors. The two main types of distortions

pertinent to our approach in 2D-3D registration are tangential + radial, and an S-shaped sigmoidal distortion (S-distortion). Subject to the earth's electro-magnetic field, an S-distortion is dependent on the pose of the image intensifier. In contrast, x-ray images from flat panel detectors exhibit little to no distortion; thus, radiographs from the Artis zeego were used directly without further rectification. Intrinsic and extrinsic parameters were taken from Siemens calibration files located on the reconstruction workstation.

To correct for image intensifier S-distortions, conventional methods have used a calibration phantom with metal BBs<sup>149,150</sup> or grooves<sup>151</sup> implanted in known geometry. These methods characterize the distortion by fitting a high-order polynomial between the observed points or line features of the grid in the x-ray image to the physical geometrical coordinates of the respective phantom grid. These phantoms are easy to use offline, but might be cumbersome when used for intra-operative distortion correction. In a more streamlined fashion, Chintalapani *et al.*<sup>152</sup> devised a novel method to perform simultaneous distortion correction and pose estimation using patient CT and characterizing variations with principal component analysis.

To calibrate and correct for distortion on images acquired using the image intensifier of an OEC 9600, we followed standard methods to fit 5<sup>th</sup> order Bernstein polynomials to known image features. This process first acquired a set of 15 images of a fluoroscopically-opaque checkerboard (20x20x20 mm squares). By modeling the X-ray system as a pinhole camera, we applied standard camera calibration (Section 3.5.1) to determine the intrinsic parameters of the 9600 C-arm, as well as polynomial corrections for tangential + radial distortion. To rectify S-distortions, we manually segmented the corners of each calibration pose, then fitted a fifth order Bernstein polynomial to the



known coordinates. For extrinsic registration, we used manual segmentation of a subset of fiducials in dual-projections of a calibration phantom (Superflab phantom, Section 4.3.4.1). Ground truth data for the OEC 9600 consists of the rectified images and projection matrices composed from these intrinsic and extrinsic calibrations. We attached two greyscale printed checkerboard markers calibrated to and visible to a Micron tracker (Claron Technology, Toronto, ON, Canada) near the detector and x-ray tube of the OEC.

### **3.6 Registration**

Registration is a key functionality for systems integrating medical imaging for augmented reality in mixed-paradigm image-guided robotics surgery. In this context, registration is the spatial alignment of different coordinate systems of various medical image data to the patient for use by the robotic system in the operating room. For our initial alignment of preoperative surgical CAD/CAM, a non-rigid transformation from CT to CBCT was followed by a rigid point-based transformation, which registers intraoperative CBCT to endoscopic video. Furthermore, 3D localization from 2D-3D registration of X-ray fluoroscopy to CBCT was explored for intraoperative updates.

#### **3.6.1 Deformable CT to CBCT Registration**

The registration of preoperative data for intraoperative use in surgery has arguably been solved for rigid anatomy. Commercially available systems exist for orthopedics as well as some neurosurgery applications, in which rigidity could be considered a reasonable assumption. However, the problem is still unsolved for deformable workspaces, such as in abdominal interventions, where non-rigid changes necessitate an update to the preoperative surgical plan. Deformations begin with perioperative setup; for

example, in trans-oral base of tongue surgery, the patient's neck is flexed, with mouth open and tongue retracted. Therefore, in order to capture patient positioning for surgery, our clinical workflow acquires a CBCT preceding *da Vinci* docking.

In a collaborative effort of the development of the described system for transoral interventions, Reaungamornrat *et al.*<sup>153</sup> developed a four step deformable registration framework for TORS to resolve the non-rigid transformation from preoperative CT to perioperative CBCT. It can be summarized as follows.

- (1) A volume of interest (e.g., tongue and hyoid bone) is segmented in both the moving image from the CT and the fixed image from the CBCT. These segmentation “masks” provide surface meshes from which two point clouds are defined.
- (2) Gaussian-mixture (GM<sup>154</sup>) registration is used to compute a rigid initial global alignment of the two point clouds (GM rigid).
- (3) A GM non-rigid registration uses a thin-plate spline approach to perform deformable alignment of the point clouds.
- (4) For both the moving and fixed mask, a distance transform<sup>155</sup> (DT) consisting of the distance of each voxel to the surface mesh is computed. A fast-symmetric-force variant of the Demons algorithm<sup>156</sup> is applied register to the two DTs.

The proposed registration was built using the Insight Segmentation and Registration Toolkit (ITK) framework. Since the Demons step operates on the DT images and not directly on image intensities, this implementation is impartial to the values and

range constraints of CT. In this work, a hybrid registration approach combines a feature-based initialization followed by a Demons refinement. Operating on distance transforms allows the combined registration module to be intensity-invariant and thereby supports registration of surgical CAD/CAM derived from other modalities, such as MRI, in addition to CT. MRI better delineates soft tissues, including the tumor and lingual nerves, whereas preoperative CT with contrast enhances vasculatures of interest, which allows for this IGRS system to take advantage of fusing multiple modalities.

### 3.6.2 Video to CBCT Rigid Registration

The endoscope, as the primary visual source in video-based robotic surgery, is the natural reference frame for our proposed image-guidance system. In fact, end effector positions are reported by the *da Vinci* API with respect to the coordinate system of the stereo endoscope.

For TORS, registration of endoscopic video with the surgical CAD/CAM, segmented from preoperative CT, is derived with the following series of transformations:

$${}^{Endoscope}T_{CT} = ({}^{Endoscope}T_{CBCT})({}^{CBCT}T_{CT}) \quad (3.5)$$

The  ${}^{CBCT}T_{CT}$  transformation is the deformable registration of preoperative CT to intraoperative CBCT, presented above in Section 3.6.1<sup>157</sup>.

Following the  ${}^{CBCT}T_{CT}$  perioperative deformation, planned data was registered to the robotic endoscope through a manual process. First, a set of fiducials from preoperative data were segmented using ITK. Through the surgeon-side using the master manipulators, an operator identified each fiducial with virtual cursors through the *cisst*

3D User Interface. The corresponding data between the two data sets provide the rigid point-based transformation,  $^{Endoscope}T_{CBCT}$ .

Preoperative CT to video registration for endoscopic skull base intervention, guided by a Medtronic StealthStation, was formulated in Equation 3.1. This manual step, a standard clinical process, used either point localization of skin surface fiducials or multi-point sampling of rigid surfaces on the patient’s face. This proprietary supported registration function was integrated into our image-guidance system.

The accuracy of the stereo video overlay for TORS was assessed by using a rigid anthropomorphic skull phantom derived from 3D rapid prototyping modeled from a cadaver CT scan<sup>158</sup>. Five fiducials on the skull surface and their corresponding point locations in a CBCT of the phantom were manually segmented. Additionally, three target fiducials embedded near the soft palate were also localized to assess the target registration error (TRE).

Using the five surface fiducials as registration markers, we overlaid the three manually segmented target fiducials onto the stereoscopic video. Next, we captured an image of the augmented scene in five different camera poses, at extensions of the workspace. Within each image, we measured the mean TRE (Projection) (i.e., the shortest distance between the overlaid CT targets and rays through the visible target seen through both cameras). From this experiment, we were able to achieve a mean TRE (Projection) of  $1.82 \pm 0.92$  mm.

### 3.6.3 Video to Critical Structures Vision-Based Rigid Registration

In the clinical workflow for non-rigid environments for TORS, the registered surgical CAD/CAM is valid only up to Step (4) on surgical approach (i.e., prior to any incision/resection). In order to update the overlay of the tumor/margin for TORS during a base of tongue oncologic dissection, our IGRS system needs to track the motion of the volume being resected. Previously, researchers have identified challenges in soft tissue surgical navigation pertaining to organ shift and tissue deformation using augmented reality in endoscopy<sup>159</sup>. As a first step toward intraoperative TORS resection updates, we used vision-based techniques to track a custom rigid fiducial attached on the surface, directly above the resection target. Assuming a constant spatial relationship within the resected volume, we updated the overlay of the tumor and margin mesh based on the rigid transformation of the tracked fiducial.

The custom fiducial was fabricated on a 3D printer and designed as a planar right isosceles triangular lattice with a hypotenuse of 10 mm in length. Each corner of the symmetric triangle was connected by an annulus, a ring with an inner radius of 1.5 mm. The green triangular frame, 1 mm in width, was embedded with three Teflon spheres (1.6 mm radius, colored white, yellow, and black), each inserted into a corner annulus. Using color thresholds, the green framework of the fiducial was first located as an initial region of interest. The negatives created by each annulus were segmented using contour detection and matched by a moving average of their previously determined color. Chromatic thresholds, updated on successful fiducial segmentations, were designed to be dynamically adaptive in order to be robust to fiducial color changes due to pollution from cautery. A rigid transformation from point-based tracking of the spheres on the

customized fiducial updated the overlay of augmenting tumor/margin meshes to follow the surgical motion of the resected volume. Experimental application of the vision-based approaching using this fiducial is described in Section 6.1.2.

### **3.7 Chapter Summary and Future Work**

This chapter describes the clinical workflow and a component-based architecture for a *robotic-endoscopic/thoracoscopic* image guidance system and an *endoscopic* image guidance system. The versatility of this design is illustrated through a reconfigurable system comprised of distinct software modules and functionalities that can be upgraded and adapted individually to a specific surgical application while leveraging the aggregate architecture. Significant components include visualization using augmented reality and a 3D user interface that support rendering and interaction of medical data from intraoperative C-arm imaging directly within the existing primary clinical view port of the *da Vinci* robotic systems. We detailed and evaluated the functional and performance requirements of the system, including calibration of the imaging and optical systems. Future work should automate all calibration steps, which are currently manual. Furthermore, our optical camera calibration is conducted at a fixed focal length, but future work should support dynamic focal length changes.

The foundation of this system builds from different classes within an existing set of libraries (*cisst*/SAW) to support each component, as well as a framework to manage their communication; however, greater emphasis is placed on the extension achieved and our contribution to the development of novel features, as described by this thesis. Innovations from the efforts of this work include a stereo alpha-blended video augmentation, features of a 3D UI that allow direct user manipulation of virtual surgical

CAD/CAM, and deformable multimodal CT to CBCT registration. Theoretically, because it is invariant to intensity, the deformable registration can process images from other modalities, such as MRI. That said, testing with clinical MRI data has not been included in this dissertation and can be an invaluable future project. In contrast to prior similar work by other groups, this image-guided *da Vinci* system is distinct in its extension of the surgeon-side console, using kinematics from the API to update stereo video augmentation, enhanced depth perception, navigation, and an integrated user interface.

Lastly, to the best of our knowledge, this system is the first of its kind to integrate guidance from an intraoperative robotic C-arm with multiple *da Vinci* systems using an adaptable design to not only address a single area of robot-assisted surgical intervention, but is also extensible and flexible enough to accommodate inter-surgical specialties. Therefore, the development of such a versatile and unique system is a meaningful contribution to research in mixed-paradigm image-guided robotic surgery system.

### 3.8 Recapitulation of Contributions

In Table 3.1 we summarize contributions made towards overcoming technical barriers of specific research problems addressed in Chapter 3.

TABLE 3.1 RESEARCH PROBLEMS (SECTION 1.4) WITH CONTRIBUTIONS TO OVERCOME  
TECHNICAL BARRIERS (SECTION 1.5)

- Problem 2: Preoperative Data and Perioperative/Intraoperative Deformation
  - Registration of preoperative image data to perioperative/intraoperative image data
    - Registration between volumetric imaging (CT, CBCT)
 

Note: The deformable CT/CBCT registration algorithm was developed by Sureerat Reaungamornrat. The author provided technical, and clinical requirements used in the conceptual design of the algorithm.
    - Registration between volumetric imaging and video
  - Intraoperative surgical motion
    - Computer vision-based resection volume tracking
- Problem 3: Effective Image Guidance
  - Adaptable software architecture and components
    - Design and implementation of a modular architecture
 

Note: The author designed and implemented all modules used in the architecture by extending cisst/SAW, the open source software infrastructure developed by Dr. Peter Kazanzides, Anton Deguet, Balazs Vagvolgyi and others.
    - Fast, reliable preoperative calibration for fluoroscopic and video image systems
  - Effective systems engineering with evaluation and validation



- Feasible and deployable clinical workflow

# 4 Integration of C-arm Fluoroscopy for Intraoperative Image-Guided Robotic Surgery

In Section 3.4, we presented the usage of CBCT/CBCTA as the fixed perioperative volumetric data anchoring the deformable registration of preoperative plans. However, as the intervention progresses, models for image guidance, especially in soft tissue target, need to be updated in order to reflect intraoperative deformation. In this chapter, we explore the advantages of C-arm fluoroscopy to provide intraoperative updates for our target *da Vinci* applications. Similar work on intraoperative surgical navigation from C-arm imaging capabilities can be divided into fluoroscopy<sup>114,160,161</sup> and CBCT<sup>126,162,163</sup>. Our goal for intraoperative navigation is to provide feedback in 3D regarding surgical progress with accuracy comparable to conventional techniques, such as optical or electromagnetic (EM) tracking.

Minimally-invasive surgery, including microscopic revisions, anterior skull base, and transoral surgery, are challenging because a variety of critical vascular and neural structures are present in a very confined space. Due to previous surgical procedures and

the destructive nature of some of the diseases, these cases have an increased risk of immediate intraoperative complications and long-term postoperative defects<sup>164,165</sup>. In such cases, different groups have applied optical and electromagnetic tracking solutions to provide navigation for neurosurgery and otologic surgery<sup>166-171</sup>. In comparison, we propose deriving 3D localization using X-ray fluoroscopy from two C-arm views to provide intraoperative image-based guidance for robotic surgery and to demonstrate the comparable accuracy of this approach with respect to conventional navigation systems.

#### **4.1 Conventional Navigation: Optical and Electromagnetic Tracking Systems**

To assist with intraoperative navigation, researchers have integrated 3D coordinate-measuring navigation devices that accurately localize calibrated tools relative to registered preoperatively generated image data sets (e.g., CT or MRI). Early tracking devices were essentially mechanical digitizers<sup>172</sup>, while optical and electromagnetic navigation technique were first explored for neurosurgery. Ecke *et al.*<sup>173</sup> conducted a prospective study comparing three optical/EM devices in both clinical (155 patients who underwent endonasal sinus surgery and 23 patients who underwent anterior or lateral skull base procedures) and laboratory conditions. Optical trackers included the EasyGuide (version 1.1 Philips, Eindhoven, The Netherlands) and VectorVision (version 3.56, BrainLAB, Munich, Germany). Electromagnetic technology included the InstaTrak (version 2.4, Vti, Woburn, MA). Laboratory testing from their results showed a much lower mean system accuracy when compared to clinical findings (e.g.,  $1.0 \pm 0.07$  mm for laboratory versus  $3.08 \pm 1.57$  mm for clinical for EasyGuide). Differences can be attributed to limitations of optical tracking systems, which may not exhibit uniform

accuracy and are difficult to optimize in clinical settings. Experimental assessment found that accuracy decreased as the surgeon proceeded further away from the centroid of the workspace. For example, with electromagnetic trackers, experiments from Ecke *et al.*<sup>173</sup> showed that the magnetic field provided by the sensors weakens at a distance of 8 cm or more from the fixed center.

Currently, most of the existing *in situ* surgical navigation systems are optical and electromagnetic (EM) solutions. For optical solutions, three types of infrared trackers are widely used in clinics: videometric, active, and passive. Videometric systems identify marker patterns on video-image sequences obtained using one or more calibrated video cameras. An example of commercially available systems that use such markers is the MicronTracker (Claron Technology Inc., Toronto, Canada), whose advantage includes a small form factor at 157 x 36 x 47 mm. In active optical trackers, light-emitting diodes (LEDs) operating in the near-infrared (IR) range are used as markers, tracked by either two planar or three linear charge-coupled device (CCD) units that form the camera module. The passive optical trackers are similar in principle to videometric systems, but they work in the near-IR range. Instead of active markers, retro-reflective spheres are illuminated by the tracking system in the near-IR spectrum. The pattern of the reflective markers, which must be unique for each tracked probe, is identified through CCD cameras. The Polaris System (Northern Digital, Inc., Waterloo, Canada) combines both active and passive infrared trackers in a single system. Compared to electromagnetic digitizers, optical digitizers offer more flexibility and are easier to use; they can also be wireless and have been quickly adopted because of their high accuracy and large field of view. However, they also require that a line of sight be maintained between the tracking

device and the instrument to be tracked, which is not always convenient and precludes the tracking of instruments inside the body. This challenge has led to the development of electromagnetic tracking systems, which have no line-of-sight requirement and are able to track instruments such as catheters and the tips of needles inside the body.

Electromagnetic digitizers localize small electromagnetic field sensors (solenoids) in a pulsed magnetic field of known geometry; thus, they superimpose a magnetic field around the surgical workspace. For this purpose, the tracking system requires a magnetic field generating source, a magnetic field-detecting sensor, and processing software. External ferromagnetic materials and additional EM fields can cause interferences, distorting the source-EM field and diminishing accuracy. However, recent developments in this technology have enabled extremely small sensor coils (less than 0.5 mm in diameter and 8 mm in length) to be embedded in surgical instruments for tracking inside the body. These systems are more robust and can better withstand such disturbances. Furthermore, special geometries have been created that can work in the operating arena. Products commercially available today that researchers can integrate into image-guided systems include the Aurora from Northern Digital, Inc., and the microBIRD from Ascension Technology Corp. (Burlington, Vermont). The potential use of electromagnetic tracking in the clinical environment and several factors to be considered have also been assessed by Yaniv *et al.*<sup>123</sup>.

Optical and EM intraoperative localizing systems provide surgeons with precise, real-time spatial registration of a patient's anatomy through the preoperative image set. A recent article comparing optical and electromagnetic tracking systems noted some additional advantages of optical tracking, such as higher accuracy<sup>174</sup>. In

otorhinolaryngology, optic and EM technologies have been found to have the widest applicability compared to other tracking principles, such as sonic or electromechanical systems. Many researchers have reported an acceptable accuracy attainable with these systems in clinic to be  $\sim 2$  mm target registration error (TRE)<sup>122,123</sup>, while others have noted higher errors within dynamic EM fields<sup>175</sup>. Although each of these conventional platforms have individual tradeoffs and potential deficiencies, they have provided clinically acceptable accuracy.

## **4.2 C-arm Fluoroscopy-Based Navigation**

C-arm X-ray radiographs are commonly used for interventional image guidance for many procedures in orthopedics and radiation therapy. X-ray fluoroscopy is a real-time and cost-effective modality for visualization of bony anatomy and surgical instruments, but image resolution and quality can vary greatly depending on the clinical target and sophistication of technology. Modern high-end CBCT-capable C-arms can now provide the surgeon with high quality three-dimensional images, depicting not only the normal anatomy and pathology, but also vascularity and function (e.g., angiography). An exploration of applying C-arm angiography has been given in Section 3.4. Interventional imaging provides real-time updates; however, for select surgical procedure, registration of navigation with preoperative data has been shown to provide sufficient guidance<sup>96,167,171,176-181</sup>.

In surgical procedures on soft tissue targets, intraoperative navigation using conventional trackers (optical, EM) and preoperative images becomes outdated once non-rigid deformation/resection occurs. C-arm x-ray fluoroscopy is widely used for localizing surgical instruments with respect to anatomical landmarks. It captures real-time

radiographic views of the surgical field, including deformation, but lacks depth information and imparts a dose of radiation to the surgeon and patient. Integration of navigation with respect to preoperative surgical CAD/CAM and intraoperative x-ray tomographic imaging can potentially address both issues by reducing the reliance on fluoroscopy and improving 3D visualization.

Various studies have proposed hybrid navigation systems that fuse an infrared optical tracking system with a mobile C-arm<sup>182-185</sup>. Drawbacks to these designs include a non-uniform range of accuracy and susceptibility to line-of-sight occlusion. Interventional imaging provides real-time updates; however, for surgeons, registration of an intraoperative image with respect to preoperative data can provide valuable information and guidance. Meanwhile, initial alignment of the C-arm with respect to the surgical target for 2D/3D imaging can be time consuming, involving multiple fluoroscopic acquisitions<sup>182</sup>. Dressel *et al.*<sup>183</sup> proposed a camera-augmented mobile C-arm (CAMC)<sup>184</sup> in which initial positioning relied on a video-based approach. To address issues regarding line-of-sight, Reaungamornrat *et al.*<sup>185</sup> created a novel tracker configuration (referred to as “Tracker-on-C”), which mounts the infrared camera directly on the gantry of a mobile C-arm. Their work also reduced fluoroscopy time during setup by using digitally reconstructed radiographs (DRRs) generated from 3D volumetric data.

A DRR is a radiographic image generated from a reconstructed CT volume from modeling the physics of a simulated X-ray system. The extraction of a 2D image from a 3D CT dataset correlates intensities absorbed along the paths of virtual X-rays casted through the volume. Thus, ray-casting is a widely-accepted technique for creating DRRs<sup>186-188</sup> and often can be optimized using parallelization from graphical processing

units. For intensity-based 2D-3D registration algorithms, a DRR must generally be generated at each iteration of the optimization algorithm, in order to calculate the similarity measure<sup>189,190</sup>. DRRs have also been used for calibration<sup>184</sup>, distortion correction<sup>152,191</sup>, and other registration<sup>185</sup> purposes. An X-ray tomographic image presents live context in a single plane of real-time instrumentation with respect to anatomical deformation. However, registration of this 2D fluoroscopic image with respect to preoperative diagnostic volumes relates a projection of the intraoperative progress back to the original CT and potentially the surgical CAD/CAM and plan. Furthermore, to provide 3D intraoperative navigation, two or more projective views from calibrated systems can resolve stereo localizations of points of interest, including tool positions. A highly accurate stereotactic localization within the context of preoperative volumetric datasets has been explored for navigational guidance in surgery<sup>192,193</sup>. Methods and implementation for the required 2D-3D registration form an active area of research with applications not only in surgery<sup>64,139</sup>, but also in interventional radiology and radiation therapy<sup>194</sup>.

Similarly, we use DRRs to align intraoperative X-ray fluoroscopy to CBCT in order to apply an image-based technique to track deformable resection targets for thoracic intervention. However, a distinction of the work presented here is the integration of the Artis zeego, a high-end robotic C-arm that is precisely installed in an operating suite (i.e. the patient table is secured to the floor at a known location relative to the C-arm base). Thus, we do not require an external tracking system, but instead take advantage of the kinematic information directly from the zeego robotic system to initialize 2D-3D registration. Our work not only uses 2D information from a single X-ray projection, but



also explores the potential for 3D intraoperative localization from X-ray projections in two C-arm views (Section 4.3.3). We set our target performance accuracy for C-arm-based 3D localization at a TRE  $\leq 2$  mm (i.e., comparable to *in situ* optical/EM navigation systems).

## 4.3 C-arm Fluoroscopy for Intraoperative Image-Guided Robotic Surgery

### 4.3.1 C-arm and Robotic Workspace Analysis

Each of our motivating clinical applications requires a different intraoperative setup; therefore, we explored and determined their individual workspace configurations for an integrated *zeego-da Vinci* intervention. In order to acquire a 3D volumetric reconstruction using the *zeego*, the C-arm needs a complete a full scan around the primary axis of the patient (i.e., left anterior oblique (LAO)/ right anterior oblique (RAO)). However, reconstruction from a half scan ( $180^\circ$  plus two fan-angles) have presented comparable results in cone-beam x-ray tomography.<sup>195-198</sup> In our analysis, we looked to determine the feasible LAO/RAO angular range for each intraoperative *zeego-da Vinci* configuration. X-ray fluoroscopic images obtained in an angular range within less than a half scan can be used in limited-angle tomographic reconstructions (e.g., digital tomosynthesis). In fact, digital tomosynthesis from cone-beam C-arms have been compared with mammography and CBCT for breast imaging purposes.<sup>199,200</sup>

In addition, from these analyses, we demonstrate the modularity of our system by adapting our robotic interface to multiple distinctive *da Vinci* models, namely the *S/Si*

and *Sp*, as well as two C-arm systems, namely the OEC 9600 and Siemens Artis zeego, each representing different ends of the technical spectrum.

Workspace analysis began with the positioning of a phantom on an operating table appropriate for a particular surgical intervention. Clinical proctors and literature reviews<sup>201,202</sup> provided references on setup guidelines, including port placement as the patient side cart (PSC) of the *da Vinci* robot is positioned and docked. To position the Artis zeego for intraoperative imaging, a collision-free path was found and executed using the commercially available syngo X workstation system (Siemens AG, Healthcare Sector, Forchheim, Germany), which remotely controls the Artis zeego. Once both the patient-side cart and C-arm were readied for robotic intervention and intraoperative imaging, respectively, we explored the remaining free space of the image system by manually rotating the C-arm, changing only in LAO/RAO. This angular range of feasible intraoperative fluoroscopic images was determined for each surgical application for potential partial-scan reconstructions and localization using multiple C-arm views.

#### ***4.3.1.1 da Vinci S/Si***

The design of the patient-side cart (PSC) of the *da Vinci S* and *Si* are comparably similar. Therefore, although the following workspace analysis was conducted only with the *Si*, these findings are also applicable to the *S* model.

##### ***a. Otolaryngology – Head and Neck Surgery (TORS, Cochlear Implant)***

For TORS, we experimentally determined that keeping the base of the PSC fixed at the initial perioperative setup still allowed for a full CBCT scan, with all robotic arms retracted (Figure 4.1 Photographs of the *da Vinci Si*-zeego workspace configuration for

transoral robotic surgery). The ability to acquire a full CBCT without repositioning the PSC, though with arms retracted, was also confirmed for cochlear implantation setup (Figure 4.2), which uses a similar patient-table approach for positioning for the PSC. The free space for these two otolaryngology – head and neck procedures therefore supports intraoperative CBCT during intervention with minimal workflow changes. If used to register preoperative plans, an intraoperative volumetric scan is a valuable means of corroborating the execution of preoperative plans.

Additionally, our workspace analysis found that during otolaryngology – head and neck intervention (using an *in vivo* porcine model for TORS and cadaveric sample for cochlear implant), the Artis zeego can still achieve a scan range of  $\sim 40^\circ$ . Thus, throughout the operation we can obtain live 2D x-ray fluoroscopy images, with angular separation at  $\sim 40^\circ$ , which offers a real-time planar projection of the scene, including tissue deformation and tool proximity. In addition, the C-arm can provide 3D localization from multiple projections with angular difference up to  $40^\circ$ . Using dual projections for 3D localization is further explored with thoracic surgery as the exemplary application below.

#### ***b. Thoracic Surgery***

For TORS resection of squamous cell carcinoma, the oral tongue is immobilized by a tongue blade and fixed to a retractor, exposing only the target half of the base of tongue. Thus, a rigid workspace assumption is valid until surgical resection. By contrast, robotic thoracic interventions require perioperative deflation of the lung in order to create a robotic workspace; therefore, the target tissue undergoes a large deformation, compared to preoperative image acquisition. Furthermore, the collapsed lung is manipulated and

deformed through palpation in order to localize the tumor. Therefore, thoracic robotic interventions experience more complex non-rigid intraoperative deformations compared to TORS, and thus served as our exemplary clinical motivation for intraoperative C-arm image-based updates for our IGRS system.

We explore using fluoroscopy for non-rigid tissue deformation during a minimally invasive thoracic intervention with a *da Vinci Si* robot, with the workspace analysis below. Furthermore, we determine the feasibility of using dual C-arm X-Ray radiographs for 3D localization as detailed in Section 4.3.4.

An anthropomorphic chest phantom was used to determine the intraoperative workspace of two distinct C-arm systems for a *da Vinci Si*-assisted thoracic intervention. The image quality of C-arm X-ray tomography varies depending on the clinical target and technology of the X-ray system. C-arms in clinical rotation range from older models using image intensifiers (II) to robotic motor-actuated flat panel detectors. In order to comprehensively explore different ends of the current spectrum of C-arm technology, we evaluated the workspace configuration of two different C-arm systems: an OEC 9600 C-arm (1998 GE OEC Medical Systems, Salt Lake City, UT) with a tri-mode 12/9/6" Image Intensifier and a Siemens Artis zeego system<sup>144</sup>, which consists of a flat panel detector (30x40 cm) synchronized with an x-ray tube mounted to a 6-degrees of freedom robot.

For thoracic intervention, the base of the PSC of the *da Vinci* is positioned at  $\sim 30^\circ$  relative to the surgical table coincident with the target patient side (Figure 4.3a). The base of the OEC C-arm was situated orthogonally to the table below the PSC, as shown in Figure 4.3b. Alternatively, the Artis zeego C-arm can approach from the head of the table to be positioned to acquire images in the same direction (Figure 4.3c). In this workspace

configuration, we were able to rotate the primary (LAO/RAO) angle of the OEC 9600 from  $0^\circ$  to  $30^\circ$ , whereas for the Artis zeego, we were able to achieve (LAO/RAO) articulation from  $+70^\circ$  to  $+115^\circ$  (RAO). This workspace evaluation showed that a *da Vinci Si* and an OEC 9600 or Artis zeego C-arm configured for thoracic interventions allows intraoperative X-ray acquisitions within a scan range of  $\leq 30^\circ$  or  $\leq 45^\circ$ , respectively.

**c.     *Gynecology***

In addition to the main clinical applications explored in this dissertation, the workspace analysis can be extended to additional surgical specialties. This supports the versatility and range of our proposed IGRS system, which, given the application-driven nature of robotic surgery, can provide advantages in clinical deployment. In gynecology, following similar steps used in the workspace analysis above, we determined that the *da Vinci Si-zeego* configuration can support an intraoperative free scan range of  $\sim 50^\circ$ .

**d.     *General***

In general surgery for hepatic intervention, our analysis did not find a compatible configuration of the zeego and *da Vinci Si* with adequate room for intraoperative x-ray tomographic scans. In order for the robotic arms of the PSC to maintain a feasible range medial above the abdomen (i.e., above the liver with the patient supine in reverse Trendelenburg position), the column of the PSC base must remain in close proximity to the operating table. This intraoperative setup leaves inadequate room for the C-arm of the zeego to rotate (LAO/RAO).

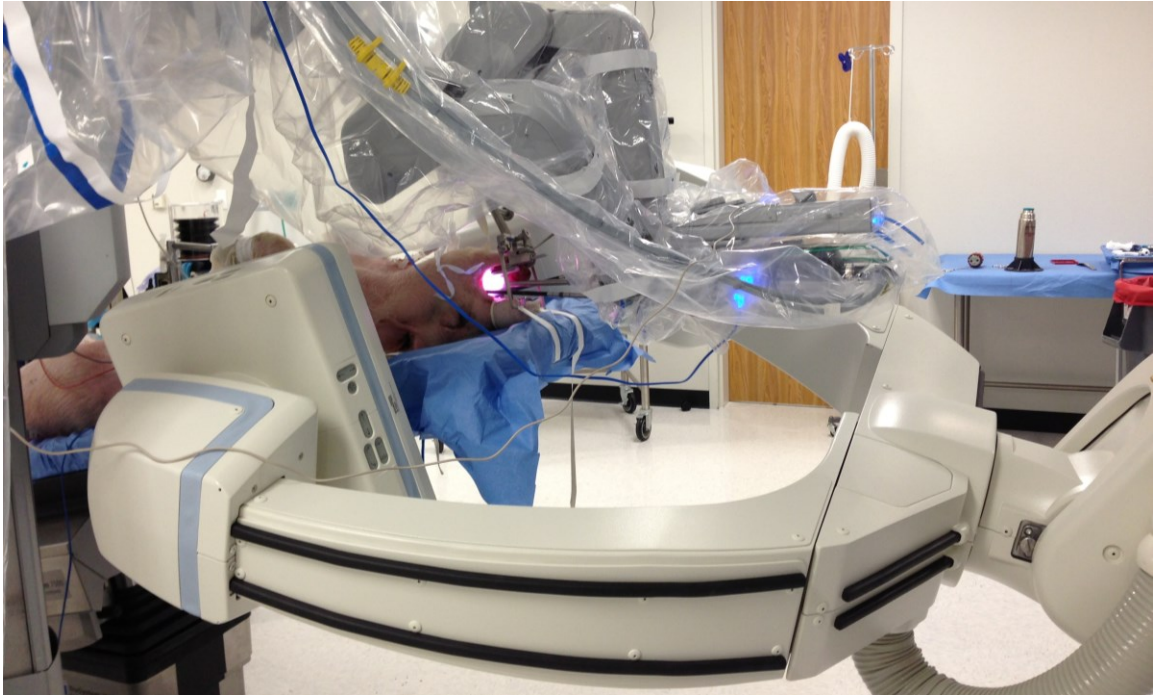


Figure 4.1 Photographs of the *da Vinci Si-zeego* workspace configuration for transoral robotic surgery.

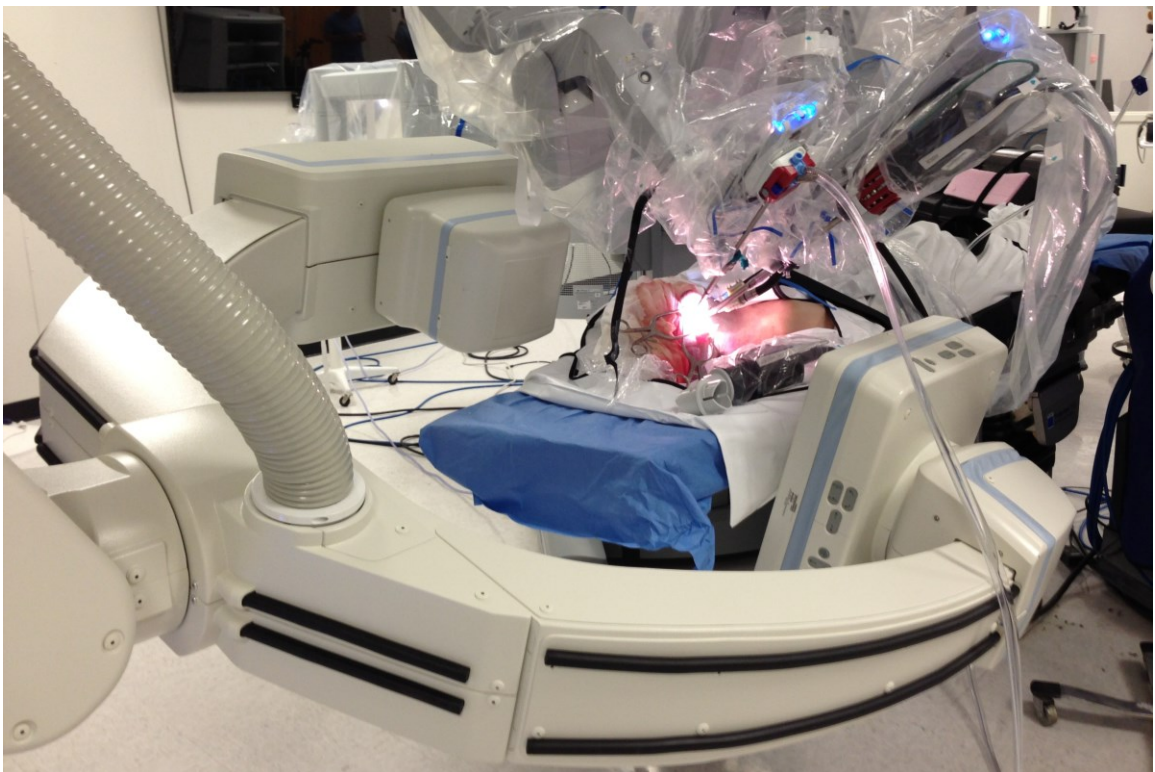




Figure 4.2 Photographs of the *da Vinci Si-zeego* workspace configuration for cochlear implant .

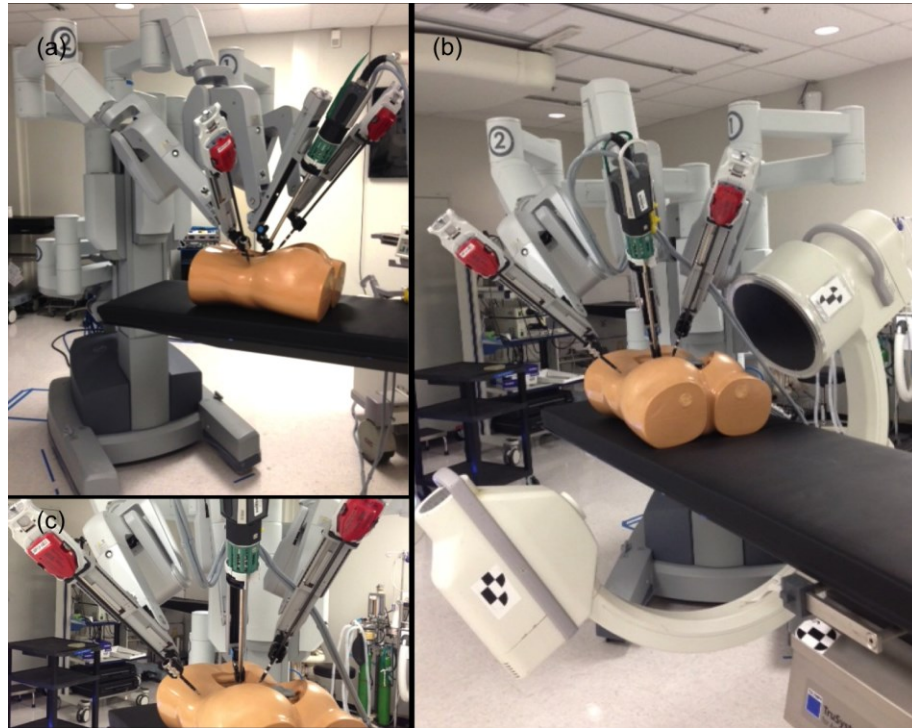


Figure 4.3 Photographs of the *daVinci-OEC* workspace configuration (a) Position of PSC at  $\sim 30^\circ$  to table (b). OEC 9600 at end of scan limit ( $30^\circ$ ) (c). Placement of *daVinci* robotic arms for thoracic intervention.

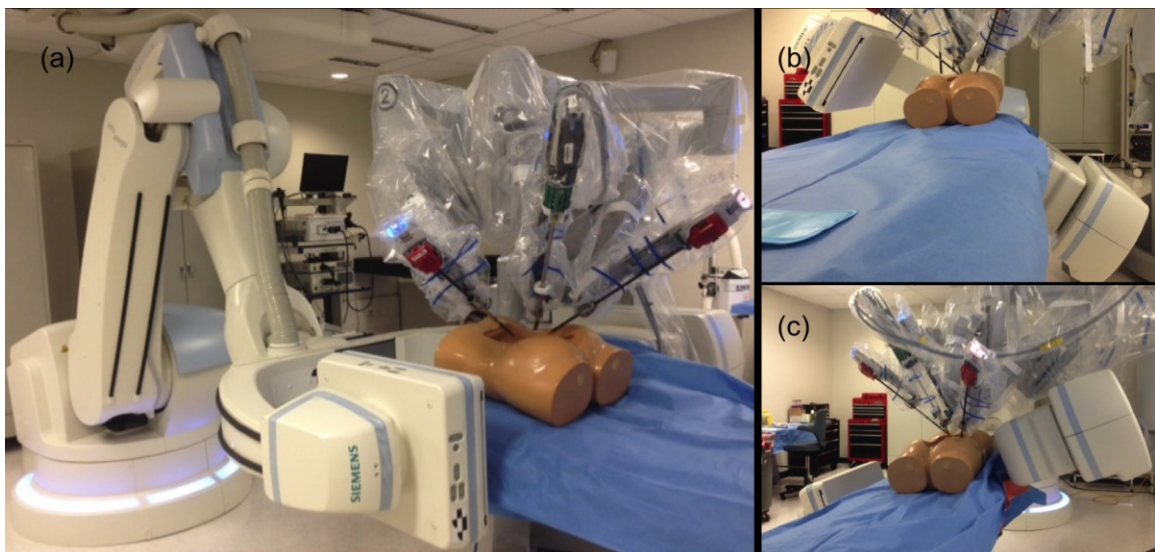


Figure 4.4 Photographs of the *da Vinci Si-zeego* workspace configuration for thoracic intervention (a) zeego positioned at table head with docked patient-side cart. zeego at the start (b) and (c) end of a  $\sim 45^\circ$   $\times$  scan range.

#### **4.3.1.2 *da Vinci Sp***

The gantry of the *da Vinci Sp* is designed with a much smaller bounding frame than that of the *S/Si* series. We further assess the flexibility of the design of our IGRS system by exploring this advantage with additional clinical applications.

##### **a. *Gynecology***

For gynecological procedures, with the patient feet-first-supine, the *da Vinci Sp* is ideal for a trans-vaginal approach (Figure 4.5). With the base of the PSC parallel to the table, the overhang boom of the *Sp* Entry Guide Manipulator (EGM) is suspended along the principal axis of the patient to direct the trocar for a natural orifice approach. In this configuration, the scan range for x-ray tomography was found to be  $\sim 100^\circ$ , which is twice the scan range of the *da Vinci Si-zeego* experiments.

##### **b. *General***

However, in general surgery for hepatic intervention, despite the streamlined design of the *Sp*, the *da Vinci Sp-zeego* workspace has a  $0^\circ$  degree scan range. The column of the PSC base must remain in close proximity to the operating table. This intraoperative setup leaves inadequate room for the C-arm of the zeego to rotate (LAO/RAO).





Figure 4.5 Photographs of the *da Vinci Sp-zeego* workspace configuration for gynecology.

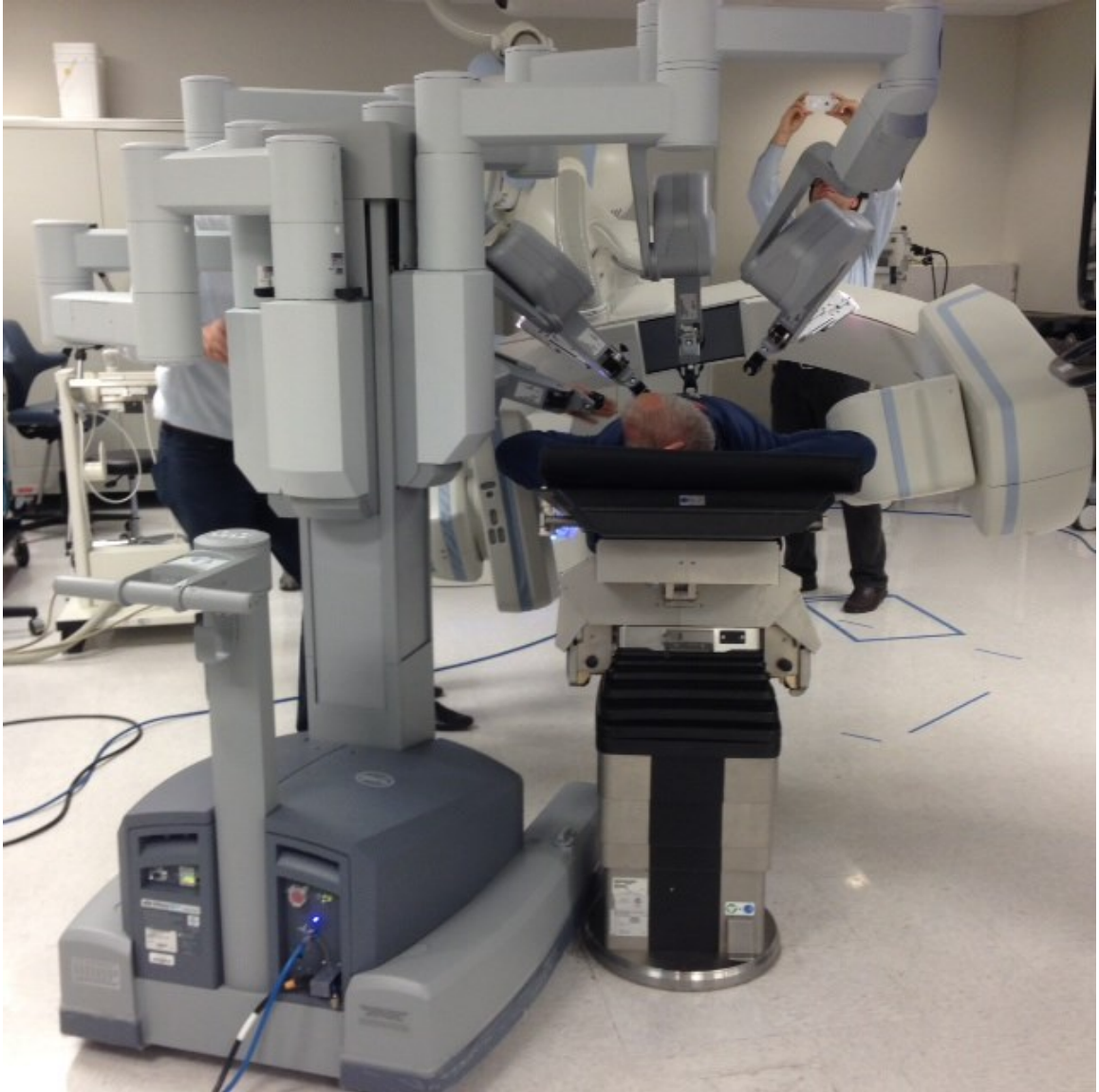


Figure 4.6 Photographs of the *da Vinci Si-zeego* workspace configuration for hepatic intervention.

#### 4.3.2 Digital Tomosynthesis

Results from the workspace analysis are summarized in Table 4.3. Although the available partial angle for all motivating clinical applications are inadequate for a full scan (i.e. theoretically requiring  $180^\circ$ + half of cone-beam angle), which is required for a 3D reconstruction using a zeego, we explore 3D localization from dual projections for

thoracic surgery. Furthermore, partial angle scans leads to digital tomosynthesis, an active topic of research for cone-beam imaging<sup>199,203-206</sup>. Digital tomosynthesis has become a promising approach for 3D volumetric imaging to detect early breast cancer<sup>200</sup>. Although conventional mammography is currently the best modality to detect early breast cancer, it is limited in that the recorded image represents the projection of a three-dimensional (3D) object onto a 2D plane. To investigate possible improvements in lesion detection accuracy with either breast tomosynthesis or digital mammography (DM), Gong *et al.*<sup>200</sup> conducted a computer simulation study that realistically modeled x-ray transport through a breast model, as well as the signal and noise propagation through a CsI-based flat-panel imager. Results indicated that for the same dose, a 5 mm lesion embedded in a structured breast phantom was detected by the volumetric breast tomosynthesis, with statistically significant higher confidence than with planar digital mammography. Results of partial reconstruction (i.e., digital tomosynthesis) of a cadaver canine thoracic phantom (described in Section 4.3.4.1) is shown in Figure 4.7 at scan angles of (a) 30° (b) 60° (c) 90° (d) 180°. Optimizing reconstruction parameters, integrating from prior images<sup>207</sup>, or providing interpolated projections from atlases<sup>208</sup> may improve these reconstructed partial scans. However, further research on this topic is beyond the scope of this dissertation.

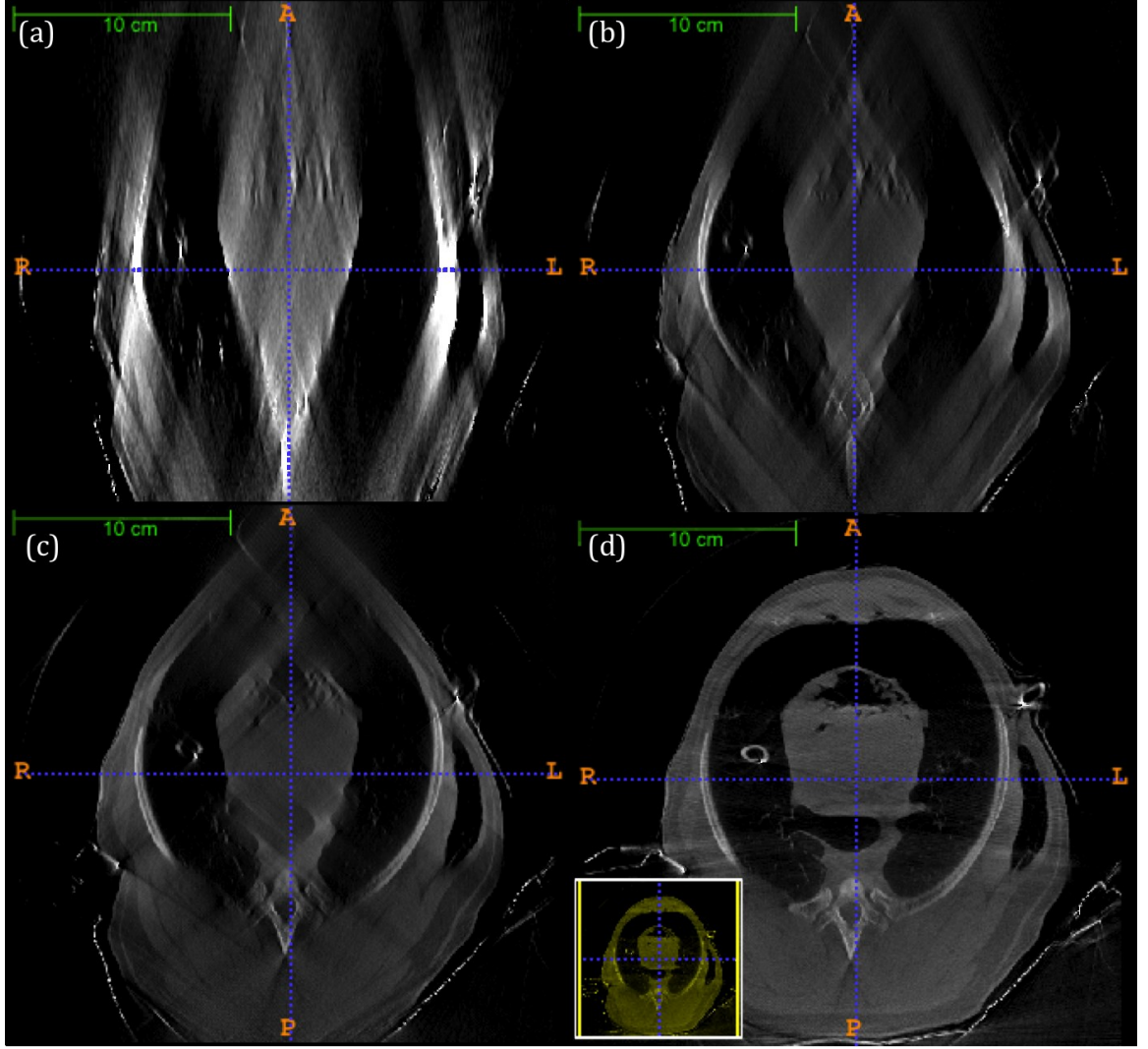


Figure 4.7 Single axial slice of canine thoracic phantom from digital tomosynthesis of (a) 30° (b) 60° (c) 90° (d) 180° scans.

### 4.3.3 Intraoperative 3D Localization from Two X-ray C-arm Views

A projective view of a single X-ray tomogram by a standard C-arm system can be modeled as a pinhole camera. Our coordinate system (i.e. the same as the one used by an Artis zeego with a patient lying head-first-supine) is as follows: x-axis (medial to left), y-axis (feet to head), and z-axis (back-to-front).

Using a pinhole model, a projection matrix,  $P \in R_{3 \times 4}$ , was used to map 3D point

locations to their corresponding 2D projection in the fluoroscopic image.

$$P = K[R|t]$$

□4.□□

The intrinsic camera parameters were captured in  $K \in R_{3 \times 3}$ , while  $R \in R_{3 \times 3}$  and  $t \in R_{3 \times 1}$  together encompassed the extrinsic camera parameters for rotation and translation, respectively. 3D point locations,  $(X, Y, Z)^T$ , were projected onto corresponding 2D image locations,  $(u, v, 1)^T$ , in a radiograph as follows

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = K[R|t] \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

□4.□□

The intrinsic camera parameter matrix

$$K = \begin{pmatrix} SDD/p & \gamma & O_x \\ 0 & SDD/p & O_y \\ 0 & 0 & 1 \end{pmatrix}$$

□4.□□

required calibration to determine source-to-detector-distance (SDD), and  $p$ , the pixel size are set as identical in  $x$  and  $y$  for the purposes of these experiments. Parameters  $(O_x, O_y)$  represent the coordinates of the isocenter in the image plane while skew,  $\gamma$ , was set to zero during our analyses in Section 4.3.4.

Extrinsic parameters captured the rigid transformation of the volumetric isocenter to the x-ray source. The primary rotation of a C-arm around the principal axis of a patient lying supine is indicated by the value of the left anterior oblique/right anterior oblique

angle (i.e., LAO/RAO). A secondary angle (cranial/caudal, i.e., CRAN/CAUD) represents how much a C-arm has been angulated toward a patient's head (cranial) or feet (caudal). For the purposes of our experiments, the 2D fluoroscopic image acquisitions utilized were composed from projective views that varied only in LAO/RAO angles. No other extrinsic angulation was performed, including CRAN/CAUD, which remained at zero.

Extrinsic rotation  $R$  can be composed as follows

$$R = R_y R_x$$

$$(4.4)$$

$$R_y = \begin{pmatrix} \cos(-\alpha) & 0 & \sin(-\alpha) \\ 0 & 1 & 0 \\ -\sin(-\alpha) & 0 & \cos(-\alpha) \end{pmatrix} R_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \quad (4.5)$$

where  $\alpha$  = LAO/RAO and  $R_x$  (rotation of 90° around x) is a standard transformation for 2D monitor displays for radiographs. The translational part of the extrinsic parameters can be defined as

$$t = (0, 0, SID)^T$$

$$(4.6)$$

where SID is the source-to-isocenter distance.

From a pair of X-ray tomographs, or dual projections, we localized the 3D position of select points of interest,  $(X, Y, Z)^T$  by triangulation through manual segmentation of their 2D image locations,  $(u, v, 1)^T$ .

$(X, Y, Z)^T$ , or 3D point location from a single projection is derived from (4.2) as

follows:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \frac{1}{\lambda} (KR)^{-1} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} (KR)^{-1} Kt \quad \square 4.7 \square$$

where  $\lambda$  represents an arbitrary scale factor. Using X-ray fluoroscopy from two C-arm views, we compute the 3D intersection point by solving a least squares formulation of (4.7). The Euclidean distance between this 3D point and its corresponding reconstructed 3D location,  $(X, Y, Z)_{CBCT}^T$  was measured as TRE<sup>120</sup>.

Two phantom models, a synthetic phantom and a canine cadaver embedded with fluoroscopic-opaque targets, were used to evaluate the workspace for the OEC 9600 and Artis zeego, respectively. For the OEC 9600, to initially align the C-arm to the patient, we use manual segmentation from five of the ten embedded PTFE spheres in dual-projections of the Superflab (synthetic oil gel) phantom. The ground truth data consists of the rectified images and projection matrices composed from these intrinsic and extrinsic calibrations. For the Artis zeego C-arm system, static geometric calibrations of given trajectories for supported isocenters are physically calibrated to compensate for these discrepancies and the flat panel detector provides distortion-free images. For our experiments, these projections are used as a ground truth starting point to which extrinsic perturbations were added. We project target points onto each fluoroscopic image using their given projection matrices. To assess the accuracy of these projections, we compared manual segmentation of the fiducials with their projected locations, which were found to be  $\leq 3.3$  pixels (Figure 4.7).

#### 4.3.4 Experimental Validation of 3D Localization from Intraoperative X-ray Fluoroscopy for Robot-Assisted Thoracic Surgery

In the following section, we apply 3D localization from two X-ray C-arm views, as discussed in Section 4.3.3, for intraoperative image-based updates for *da Vinci* robot-assisted thoracic surgery. We present an experimental evaluation of dual-projection geometries with an intensity-based 2D-3D registration of intraoperative radiographs to cone-beam computed tomography (CBCT). Using phantom models, we determine the feasible range of x-ray projections achievable by a C-arm positioned around a *da Vinci Si*, configured for robotic lobectomy. Experiments were conducted on synthetic and animal phantoms imaged with an OEC 9600 and a Siemens Artis zeego, representing different C-arm systems currently available in clinical use. The range of angular difference,  $\Delta\theta$ , of dual C-arm projections varied from  $\sim 0^\circ - 90^\circ$ , while extrinsic and intrinsic geometric parameters were tuned to the achievable intraoperative workspace of each C-arm and *da Vinci* setup. Results show that using either an optically tracked OEC 9600 or a Siemens Artis zeego, an angular difference  $\Delta\theta: \sim 30^\circ$  achieves  $\text{TRE}_{\text{mean}} \leq 2.5 \text{ mm}$ , and  $\text{TRE}_{\text{mean}} \leq 2.0 \text{ mm}$ , respectively (i.e., comparable to standard clinical intraoperative navigation systems).

In order to determine the range of feasible intraoperative X-ray projections available during a C-arm-guided *da Vinci Si* thoracic intervention, we conducted a workspace evaluation for each C-arm system, detailed in Section 4.3.4.2. Furthermore we present experimental variation of dual-projection and geometric uncertainty modeled after the constraints derived from workspace configuration experiments. Two phantom models, a synthetic phantom and a canine cadaver embedded with fluoroscopic-opaque



targets as described in Section 4.3.4.1, are used to evaluate the workspace for the OEC 9600 and Artis zeego, respectively.

#### ***4.3.4.1 Thoracic Phantom Image Acquisition***

Evaluation of each C-arm was conducted on different phantoms. For the OEC 9600 we attached a block of *Superflab* (~50 mm x ~130 mm x ~20 mm) embedded with 10 polytetrafluoroethylene (PTFE) fiducial spheres (1.6 mm diameter), to the synthetic spine of the torso phantom. Half of the fiducials were used for registration, while the other 5 were used for evaluation. CBCT images of this phantom was acquired using Siemens syngo DynaCT, 90 kVp, 290 mA, (0.48x0.48x0.48 mm<sup>3</sup> voxel size). We then collected a series of radiographs by rotating the 9600 C-arm from 0° to 30° while recording the transformation of the optical markers attached to the source and detector.

A canine cadaver phantom was used to assess image registration with the Artis zeego. To create a mock tumor target, a urethane medium durometer spherical medical balloon (10 mm diameter) was filled with a mixture of 0.5 ml rigid polyurethane foam (FOAM-IT®) and 0.25 ml of acrylic paint. The Siemens syngo iGuide system was used to plan the placement of the tumor and four peri-tumor metal fiducials (52100 Chromium 1 mm diameter spheres). The mock tumor was then placed in the phantom's lung, right lower lobe, using an FEP I.V. catheter (Abbocath®-T 14G x140 mm) and confirmed with real-time fluoroscopy. Another volumetric data set capturing the tumor and fiducials in the inflated lung was acquired using the same CBCT protocol as above.

#### ***4.3.4.2 Image Analysis for Dual X-ray Projection Geometries***

##### ***a. OEC 9600***

With the ubiquity of conventional optical and electromagnetic tracking systems in image-guided surgery, for the OEC 9600 system we experimented with two configurations: the C-arm with and without an optical tracker. The primary (LAO/RAO) rotation of an OEC 9600 C-arm is controlled with a passive lock mechanism indicated with visual markers drawn at  $2.5^\circ$  increments; therefore, we model the range of uncertainty in extrinsic parameters of the OEC 9600 at  $\pm 3$  degrees. We estimate translational uncertainty at  $\pm 10$  mm, comparable to the model for zeego discussed below. Using the *Superflab* dataset for experiment #1 (Table 4.1), we added extrinsic perturbations, i.e. ( $\Delta R: 3^\circ, \Delta T: 10$  mm), and measured the effect of simulated systematic uncertainty on the target locations. We then registered the projection using the 2D-3D algorithm and measured the TRE for 5 target spheres (non-registration fiducials) using 20 runs (i.e., 100 fiducial point measurements). To model an alternative C-arm setup, an OEC with optical tracking, we modified the range of OEC 9600 extrinsic uncertainty at  $\pm 2$  mm and  $\pm 1$  degree, comparable to accuracies of optical trackers as reported in the clinic. We repeated the same extrinsic uncertainty evaluation with perturbations in the range ( $\Delta R: 1^\circ, \Delta T: 2$  mm) for experiment #2.

To evaluate OEC 9600 extrinsics, the mean, median, and max TRE for the dual projection experiments, obtained at a  $30^\circ$  angular separation, are summarized in Table 4.2. Three sets of images are processed including the (A) original fluoroscopic images, (B) images after radial and tangential rectification and (C) images after radial, tangential and S-distortion correction. To emulate standard OEC 9600 extrinsic uncertainty, perturbations of ( $\Delta R: 3^\circ, \Delta T: 10$  mm) for experiment #1 showed a mean TRE at 46.0 mm, 44.7 mm, and 9.6 mm, for (A), (B), and (C), respectively. Perturbations

of ( $\Delta R: 1^\circ, \Delta T: 2 \text{ mm}$ ), representing an OEC 9600 with Micron Tracker in experiment #2, produced mean TREs for (A), (B), (C) at 31.1 mm, 26.8 mm, and 2.4 mm, respectively.

### ***b. Artis zeego***

The CBCT images of the canine phantom were reconstructed by a Siemens workstation using 496 fluoroscopic projections acquired in an  $\sim 180^\circ$  (LAO/RAO) trajectory. Using the same projection matrices from the Siemens reconstruction, we created a set of corresponding synthetic X-rays (i.e., DRR). From these two datasets, we chose combinations of pairs of images, separated by varied angular difference in the range of  $\Delta\theta: 1^\circ - 90^\circ$ , and triangulated through the four peri-tumor fiducials to measure TRE. The Artis zeego monitor display shows rotational angles in degrees and translations in centimeters. Therefore we model the range of Artis zeego extrinsic uncertainty to be  $\pm 10 \text{ mm}$  and  $\pm 1 \text{ degree}$ . Using the canine dataset, we added extrinsic perturbations in the range ( $\Delta R: 1^\circ, \Delta T: 1 - 10 \text{ mm}$ ) and measured the effect of simulated systematic uncertainty on the target locations. On both synthetic DRR (experiment #4) and zeego datasets (experiment #3) after imposing extrinsic error, we then registered the projection using the 2D-3D algorithm and measured the TRE for 25 runs (i.e., 100 fiducial point measurements). Using the *canine* dataset, we conducted additional experiments to evaluate individual rotation versus translation effects by extending extrinsic perturbations from ( $\Delta R: 1 - 10^\circ, \Delta T: 0 \text{ mm}$ ) and ( $\Delta R: 0^\circ, \Delta T: 5 - 50 \text{ mm}$ ) for experiments 5 and 6, respectively. To assess robustness against intrinsic uncertainty, we conducted additional experiments to evaluate individual focal length and optical center requirements by extending intrinsic perturbations from ( $\Delta F: 1 - 10 \text{ mm}$ ) and ( $\Delta O: 1 - 30 \text{ pix}$ ), in

experiments 7 and 8, respectively.

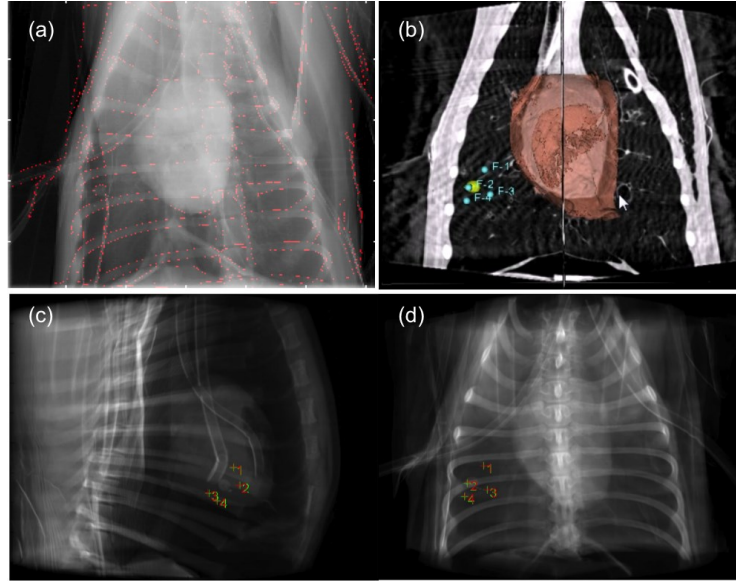


Figure 4.8 Radiographs during 2D-3D registration and DRR for ground truth assessment.

(a) Anterior-posterior radiograph with gradient (red) of registered DRR. (b) Single coronal slice of reconstructed canine thorax with segmented heart (orange), mock tumor (yellow) and numbered peri-tumor fiducials (blue). (c) Sagittal and (d) coronal fluoroscopic image with forward projections of segmented 3D targets.

Table 4.1 Experimental Protocol for 3D Localization Experiments

Exp #	C-arm system	Image	Phantom	$\Delta\theta$ [°]	$\Delta F$ [mm]	$\Delta O$ [pix]	$\Delta R$ [°]	$\Delta T$ [mm]
1	OEC 9600	radiograph	Superflab+spine	30	0	0	3	10
2	OEC 9600 w/Micron	radiograph	Superflab+spine	30	0	0	1	2
3	zeego	radiograph	canine	1-90	0	0	1	1-10
4	zeego	DRR	canine	1-90	0	0	1	1-10
5	zeego	radiograph	canine	1-90	0	0	1-10	0
6	zeego	radiograph	canine	1-90	0	0	0	5-50

7	zeego	radiograph	canine	1-90	1-10	0	0	0
8	zeego	radiograph	canine	1-90	0	1-30	0	0

Mean, median and max TRE for the dual projection experiments are summarized in the left column of Figure 4.8. Perturbations of experiment #3 ( $\Delta R: 1^\circ, \Delta T: 10 \text{ mm}$ ), simulate the maximum extrinsic uncertainty for the Artis zeego in rotation and translation. Applying this range of error generate  $TRE_{\max} \leq 14 \text{ mm}$  at angular differences,  $\Delta\theta: 30^\circ - 45^\circ$ . Using 2D-3D registration following such extrinsic perturbations, the *DRR* dataset achieved  $\leq 0.5 \text{ mm}$  for all TRE with angular differences  $\geq 15^\circ$  (experiment #4, box plot in Figure 4.8a). However, repeating the same experiment for the *Fluoro* radiographic images resulted in a  $TRE_{\text{mean}} \leq 2.0 \text{ mm}$  only for angles  $\Delta\theta: 30^\circ - 45^\circ$  (box plot in Figure 4.9b).

We investigated individual effects of rotation and translation error by extending the perturbation range for isolated  $\Delta R$  and  $\Delta T$ . For experiment #5, using the *Fluoro* dataset and introducing only rotation error ( $\Delta R: 1 - 10^\circ, \Delta T: 0 \text{ mm}$ ) we generate the box plot series shown in Figure 4.10. For the target angular difference range,  $\Delta\theta: 30^\circ - 45^\circ$ , the current 2D-3D algorithm can still achieve a  $TRE_{\text{mean}} \leq 2.0 \text{ mm}$  for rotational disturbance  $\leq 4^\circ$ . Similarly, for translation as shown in experiment #6, box plots of  $TRE_{\text{mean}}$  from perturbations of ( $\Delta R: 0^\circ, \Delta T: 5 - 50 \text{ mm}$ ) are shown in Figure 4.9. For the target range of angular difference,  $\Delta\theta: 30^\circ - 45^\circ$ , the current 2D-3D algorithm achieve a  $TRE_{\text{mean}} \leq 2.0 \text{ mm}$  for translational disturbances  $\leq 25 \text{ mm}$ .

Additionally, we investigate the effects of intrinsic uncertainty by perturbing

$\Delta F$  and  $\Delta O$ . Using the *canine* dataset and introducing focal length errors for experiment #7, i.e. ( $\Delta F: 1 - 10 \text{ mm}$ ) produce the box plot series shown in Figure 4.10. For the target angular difference range,  $\Delta\theta: 30^\circ - 45^\circ$ , the current 2D-3D algorithm can still achieve a  $\text{TRE}_{\text{mean}} \leq 2.0 \text{ mm}$  for focal length disturbance  $\leq 10.0 \text{ mm}$ . Similarly, for optical center, box plots of  $\text{TRE}_{\text{mean}}$  from perturbations of ( $\Delta O: 1 - 30 \text{ pixels}$ ) are shown in Figure 4.10. For the target range of angular difference,  $\Delta\theta: 30^\circ - 45^\circ$ , the current 2D-3D algorithm achieve a  $\text{TRE}_{\text{mean}} \leq 2.0 \text{ mm}$  for optical center disturbances  $\leq 10 \text{ pixels}$ .

TABLE 4.2 OEC 9600 EXTRINSIC EXPERIMENTAL RESULTS

OEC	$\Delta\theta_{\text{30}}^{\text{30}} \Delta R_{\text{2-3}}^{\text{30}} \Delta t_{\text{2-3}}^{\text{30}}$	$\Delta\theta_{\text{30}}^{\text{30}} \Delta R_{\text{2-3}}^{\text{30}} \Delta t_{\text{2-3}}^{\text{30}}$	$\Delta\theta_{\text{30}}^{\text{30}} \Delta R_{\text{2-3}}^{\text{30}} \Delta t_{\text{2-3}}^{\text{30}}$
<b>Rectification</b>	<b><math>TRE_{\text{mean}}</math></b>	<b><math>TRE_{\text{median}}</math></b>	<b><math>TRE_{\text{max}}</math></b>
None	43.4	43.5	60.9
Radial,Tangential	37.2	37.6	41.6
S,Radial,Tangential	8.7	8.8	18.7

OEC	$\Delta\theta_{\text{30}}^{\text{30}} \Delta R_{\text{2-3}}^{\text{30}} \Delta t_{\text{2-3}}^{\text{30}}$	$\Delta\theta_{\text{30}}^{\text{30}} \Delta R_{\text{2-3}}^{\text{30}} \Delta t_{\text{2-3}}^{\text{30}}$	$\Delta\theta_{\text{30}}^{\text{30}} \Delta R_{\text{2-3}}^{\text{30}} \Delta t_{\text{2-3}}^{\text{30}}$
<b>Rectification</b>	<b><math>TRE_{\text{mean}}</math></b>	<b><math>TRE_{\text{median}}</math></b>	<b><math>TRE_{\text{max}}</math></b>
None	27.7	27.5	31.2
Radial,Tangential	4.8	4.8	6.2
S,Radial,Tangential	3.7	3.7	6.9

OEC	$\Delta\theta_{\text{30}}^{\text{30}} \Delta R_{\text{2-3}}^{\text{30}} \Delta t_{\text{2-3}}^{\text{30}}$	$\Delta\theta_{\text{30}}^{\text{30}} \Delta R_{\text{2-3}}^{\text{30}} \Delta t_{\text{2-3}}^{\text{30}}$	$\Delta\theta_{\text{30}}^{\text{30}} \Delta R_{\text{2-3}}^{\text{30}} \Delta t_{\text{2-3}}^{\text{30}}$
<b>Rectification</b>	<b><math>TRE_{\text{mean}}</math></b>	<b><math>TRE_{\text{median}}</math></b>	<b><math>TRE_{\text{max}}</math></b>
None	46.0	46.0	56.4
Radial,Tangential	44.7	45.2	46.7
S,Radial,Tangential	9.6	9.7	17.8

OECw/Micron	$\Delta\theta_{\text{30}}^{\text{30}} \Delta R_{\text{2-3}}^{\text{30}} \Delta t_{\text{2-3}}^{\text{30}}$	$\Delta\theta_{\text{30}}^{\text{30}} \Delta R_{\text{2-3}}^{\text{30}} \Delta t_{\text{2-3}}^{\text{30}}$	$\Delta\theta_{\text{30}}^{\text{30}} \Delta R_{\text{2-3}}^{\text{30}} \Delta t_{\text{2-3}}^{\text{30}}$
<b>Rectification</b>	<b><math>TRE_{\text{mean}}</math></b>	<b><math>TRE_{\text{median}}</math></b>	<b><math>TRE_{\text{max}}</math></b>
None	29.2	29.2	36.1
Radial,Tangential	25.3	25.6	32.7
S,Radial,Tangential	2.5	2.5	7.1

OECw/Micron	$\Delta\theta_{\text{30}}^{\text{30}} \Delta R_{\text{2-3}}^{\text{30}} \Delta t_{\text{2-3}}^{\text{30}}$	$\Delta\theta_{\text{30}}^{\text{30}} \Delta R_{\text{2-3}}^{\text{30}} \Delta t_{\text{2-3}}^{\text{30}}$	$\Delta\theta_{\text{30}}^{\text{30}} \Delta R_{\text{2-3}}^{\text{30}} \Delta t_{\text{2-3}}^{\text{30}}$
<b>Rectification</b>	<b><math>TRE_{\text{mean}}</math></b>	<b><math>TRE_{\text{median}}</math></b>	<b><math>TRE_{\text{max}}</math></b>
None	21.8	21.6	24.1
Radial,Tangential	3.6	3.5	4.5
S,Radial,Tangential	2.1	2.0	4.2

OECw/Micron	$\Delta\theta_{\text{30}}^{\text{30}} \Delta R_{\text{2-3}}^{\text{30}} \Delta t_{\text{2-3}}^{\text{30}}$	$\Delta\theta_{\text{30}}^{\text{30}} \Delta R_{\text{2-3}}^{\text{30}} \Delta t_{\text{2-3}}^{\text{30}}$	$\Delta\theta_{\text{30}}^{\text{30}} \Delta R_{\text{2-3}}^{\text{30}} \Delta t_{\text{2-3}}^{\text{30}}$
<b>Rectification</b>	<b><math>TRE_{\text{mean}}</math></b>	<b><math>TRE_{\text{median}}</math></b>	<b><math>TRE_{\text{max}}</math></b>
None	31.1	31.0	34.9
Radial,Tangential	26.8	27.1	34.0
S,Radial,Tangential	2.4	2.5	9.8

<b>Perturb</b>	<b><math>\Delta t = 10</math></b>	<b><math>\Delta R = 1</math></b>	<b>n=100</b>
<b>Angle</b>	<b><math>TRE_{mean}</math></b>	<b><math>TRE_{median}</math></b>	<b><math>TRE_{max}</math></b>
15	41.8	41.7	43.1
30	12.7	12.6	13.3
45	5.0	5.1	5.2

<b>DRR</b>	<b><math>\Delta t = 10</math></b>	<b><math>\Delta R = 1</math></b>	<b>n=100</b>
<b>Angle</b>	<b><math>TRE_{mean}</math></b>	<b><math>TRE_{median}</math></b>	<b><math>TRE_{max}</math></b>
15	0.1	0.1	0.4
30	0.1	0.1	0.2
45	0.0	0.0	0.1

<b>Fluoro</b>	<b><math>\Delta t = 10</math></b>	<b><math>\Delta R_z = 1</math></b>	<b>n=100</b>
<b>Angle</b>	<b><math>TRE_{mean}</math></b>	<b><math>TRE_{median}</math></b>	<b><math>TRE_{max}</math></b>
15	3.1	3.1	3.4
30	2.0	2.0	2.2
45	2.0	2.0	2.2

\*TRE [mm]

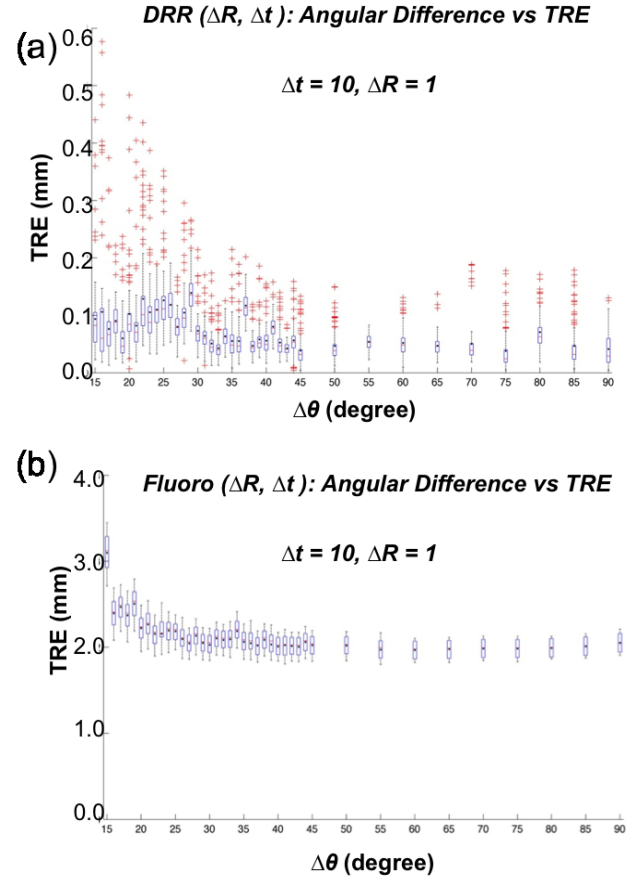
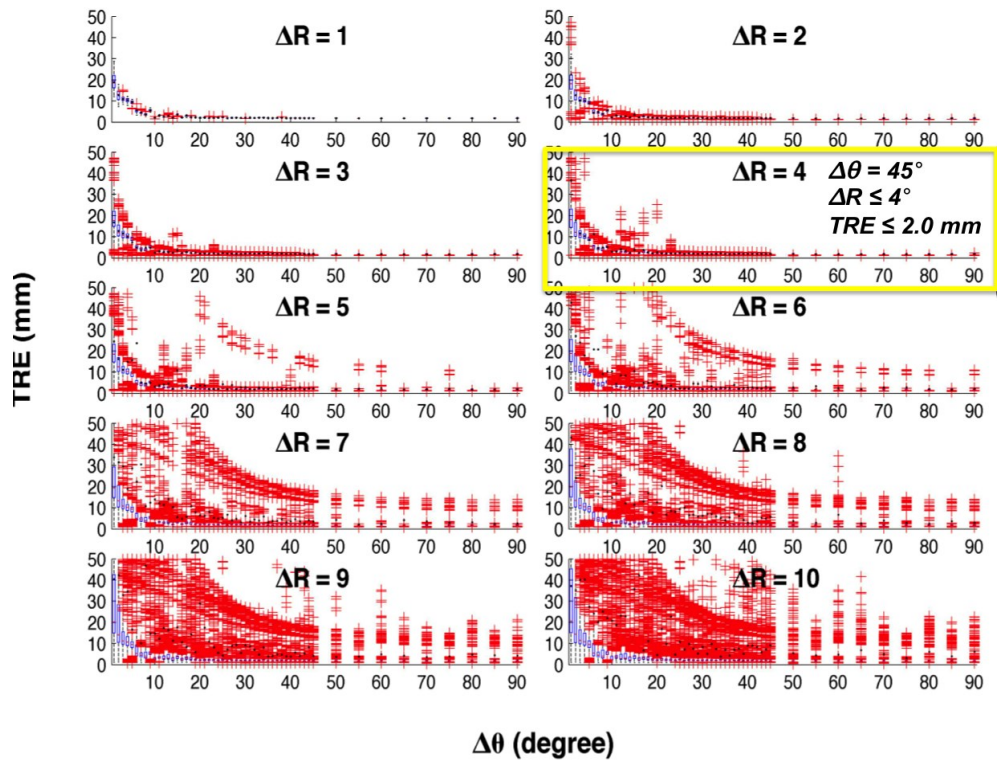


Figure 4.9 Results from dual projection experiments focused on Artis zeego extrinsic uncertainties. Table (left, top) TRE from perturbing fluoroscopic data. Table (left, middle). TRE after 2D-3D registration on perturbed DRR data. Table (left, bottom) TRE after after 2D-3D registration on perturbed fluoroscopic data. (a). Box plot of the zeego extrinsic uncertainty on DRR TRE (experiment #4). (b). Box plot of the zeego extrinsic uncertainty on fluoroscopic TRE (experiment #3).



**$\Delta R$ : Angular Difference vs TRE**



**$\Delta t$ : Angular Difference vs TRE**

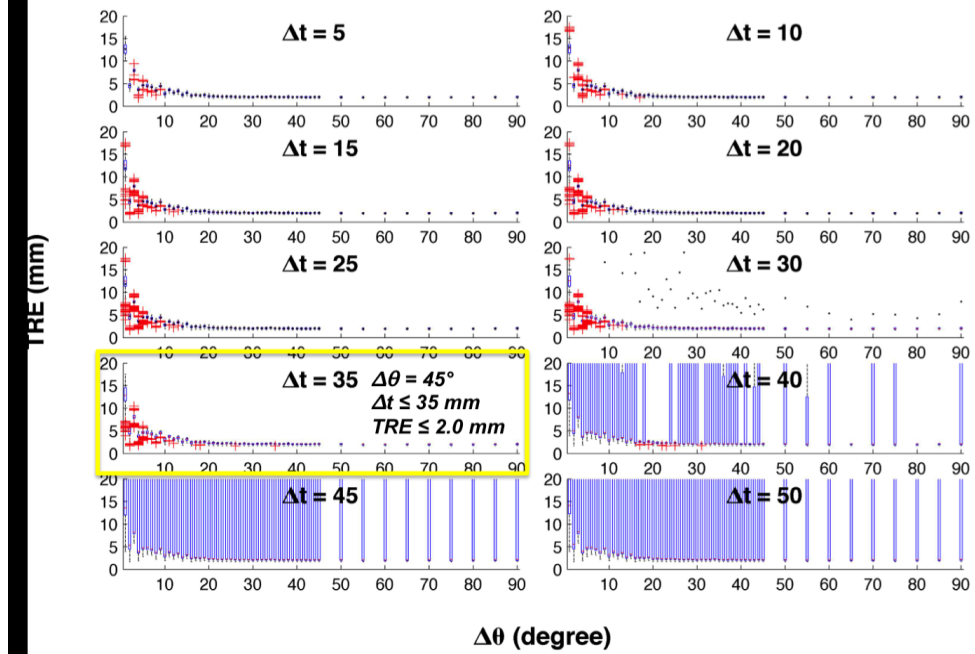
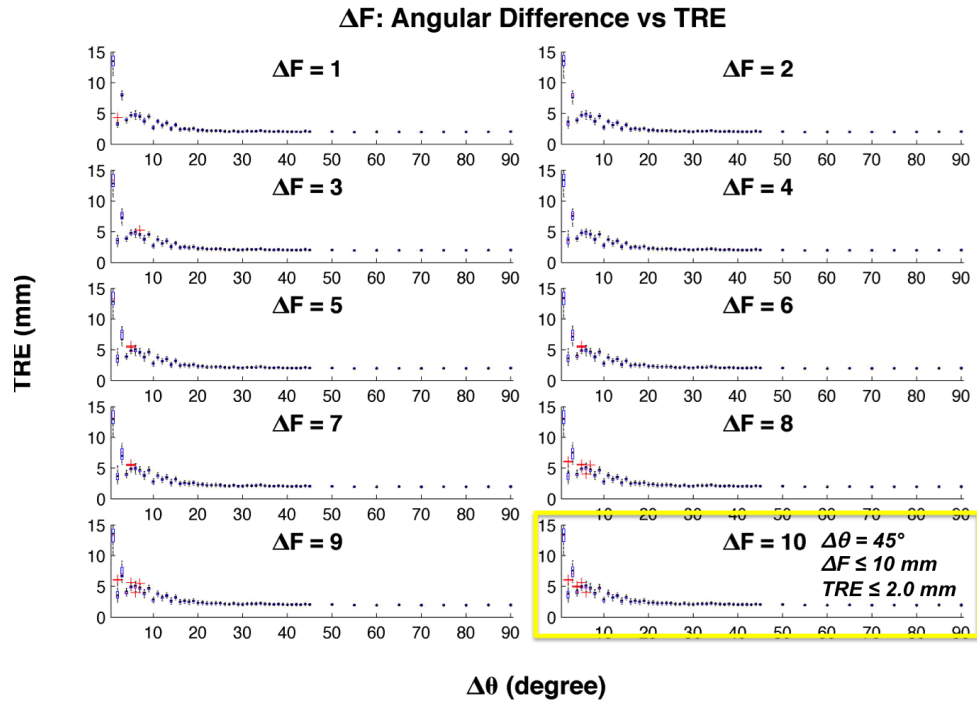


Figure 4.10 (Top: Experiment #5) Box plot of the effect of angular difference on TRE with various rotational perturbations. (Bottom: Experiment #6) Box plot of the effect of angular difference on TRE with various translational perturbations.



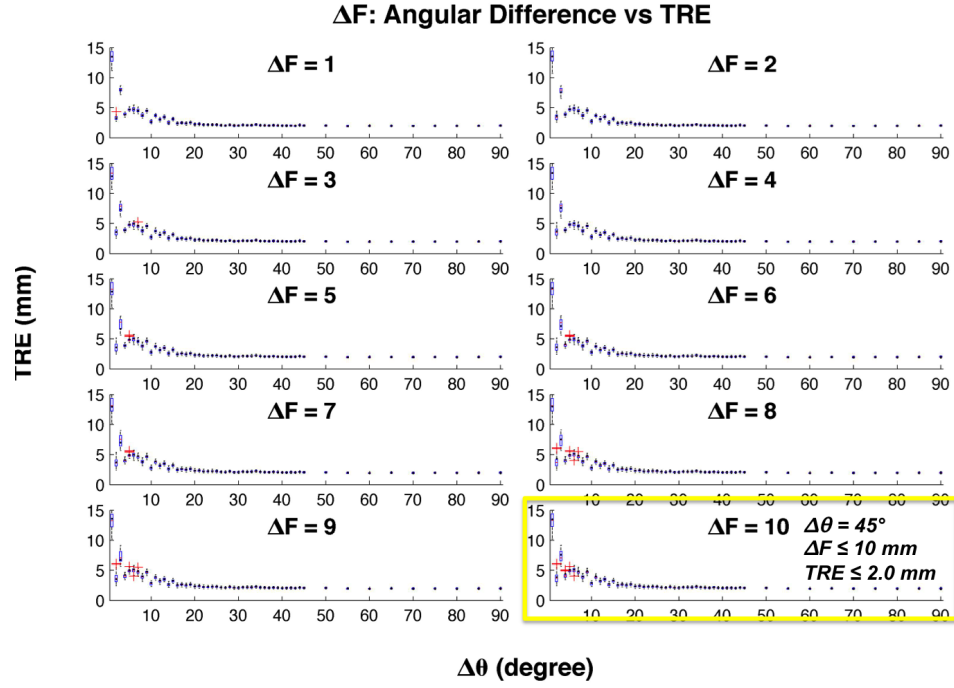


Figure 4.11 (Top: Experiment #7) Box plot of the effect of angular difference on TRE with various focal length perturbations. (Bottom: Experiment #8) Box plot of the effect of angular difference on TRE with various optical center perturbations.

#### 4.3.4.3 Discussion

Work presented in Section 4.3.4 has experimentally shown the feasibility of dual-projective radiographs to provide adequate 3D localization for minimally-invasive robotic thoracic surgery. Using phantom models, we determine the achievable range of X-ray projections by two different C-arm systems positioned around a *da Vinci Si*, as configured for robotic thoracic intervention. Currently for operable lung cancer, a lobectomy with systematic lymphadenectomy is standard of care; however, for smaller tumors, a less extended anatomical resection such as a segmentectomy might become the appropriate surgical treatment in the near future<sup>31</sup>. For these interventions, a desirable TRE of  $\leq 2$  mm (i.e., comparable to current intraoperative navigation systems) was

achieved using the proposed method of 2D-3D registration. Nominal 2D-3D registration parameters were derived from previous work<sup>209</sup> and limitations of extrinsic parameters of the *zeego* and *da Vinci* workspace used in these experiments are determined with clinically relevant in-vivo phantoms and workflows. Other groups have explored the potential of 3D localization using two C-arm views<sup>192,193</sup>, including with target application in image-guided surgery<sup>210</sup>. However, our work validates clinically relevant workspace scenarios through *in vivo* experiments and directly explores the feasible intraoperative configurations of two clinically available robotic systems.

2D-3D registration of X-ray image intensifier images varies greatly depending on the level of distortion correction applied. Raw X-ray images from the OEC 9600 were particularly susceptible to rotational disturbances likely due to a smaller field-of-view, lower contrast, as well as factors from distortion. A minimum mean TRE of 2.4 mm was achieved only with radial, tangential, and S-distortion correction using an OEC 9600 with an optical tracker. While work presented requires a one-time geometric C-arm calibration for both C-arm systems, other groups have debated its necessity<sup>211</sup>. Calibration for S-distortion at each possible pose is an unrealistic requirement for clinical deployment. To counter these issues, other groups have developed reliable statistical characterization of C-arm distortion from sparse calibration<sup>191</sup>.

Compared to synthetic DRRs, experiments using X-ray data from the Artis *zeego* (flat-panel detector) result in a much higher range of errors. Differences between these data sets can be attributed to lower signal-to-noise ratios and limitations of geometric calibrations for real fluoroscopy. Although radio-opaque peri-tumor fiducials can be inserted using the iGuide system of the *zeego* C-arm, fiducial markers placed during

preoperative biopsy may be more clinically feasible with regards to minimizing intraoperative workflow changes. This alternative workflow is more applicable to other generic C-arm systems. Conversely, though a single point fiducial is adequate for wedge resections, placing multiple peri-tumor markers may improve TRE for vascular dissection, which is required in lobectomies and segmentectomies. Furthermore, additional fiducials would better constrain tracking of the orientation of the tumor, which has applications in real-time video augmentation, as well as deformable registration.

Throughout the experiments presented, the 2D-3D registration of a single X-ray image was completed in  $\leq 8$  *seconds*, using a CUDA implementation on an NVIDIA Titan graphics card (NVIDIA, Inc., Santa Clara, CA). We expect to achieve 3D localization  $\leq 1$  *minute*, with the following required steps completed within the indicated time:

- a). Acquire X-ray #1 (1 second)
- b). 2D-3D registration of #1 (8 seconds)
- c). Segmentation of points of interest in #1 (5 seconds)
- d). Rotate zeego to  $\sim 45^\circ$  (5 seconds)
- e). Repeat a-c for X-ray #2 (14 seconds)
- f). Triangulate points from #1 and #2 (1 second)

Next steps can potentially look into clinical evaluation of the proposed workflow. We anticipate that 3D localization performed in under 1 minute will be adequately fast for an initial evaluation. Further tuning of the 2D-3D registration parameters, as well as hardware upgrades and automatic segmentation software, can accelerate this step even more. Other proposed requirements added to the standard workflow increase

intraoperative setup by 6 minutes + 5-15 minutes for each intraoperatively placed fiducial. However, the guidance provided can arguably reduce time required to locate targets of interest and delineate resection boundaries, especially in complicated cases.

#### **4.4 Chapter Summary and Future Work**

This chapter presented the integration of two clinically available C-arms (e.g. OEC 9600 and Siemens Artis zeego) with various models of the *da Vinci* surgical robotic system (e.g., *da Vinci S*, *Si* and *da Vinci Sp*). Therefore, in addition to using perioperative CBCT/CBCTA, our proposed design leverage navigational updates from 3D localization using X-ray fluoroscopy from two C-arm views. *In vivo* experimentation with the OEC 9600 and other more economical C-arms would demonstrate the breadth of the approach and should be considered as next steps. Other groups have explored the potential of 3D localization using two C-arm views<sup>192,193</sup>, including target applications in image-guided surgery<sup>210</sup>; however, to the best of our knowledge, our work is the first integration of a high-end robotic C-arm for multiple *da Vinci*-assisted interventions.

We contribute a workspace analysis by exploring the free space configurations of C-arms during various *da Vinci*-assisted clinical scenarios. Key results are summarized in Table 4.3. Future efforts should further explore image guidance from interventional fluoroscopy and possibilities for digital tomosynthesis from partial scans. For otolaryngology – head and neck procedures, we concluded that an intraoperative CBCT is achievable throughout the intervention without relocating the base of the patient-side cart, though with retraction of the robotic arms. The feasible range of real-time X-ray tomography for intraoperative 3D localization was determined for our target clinical *da Vinci* applications using a *da Vinci Si* with an Artis zeego. Exploratory experiments

validated the feasibility and significance of intraoperative 3D localization for thoracic interventions, where preoperative data for soft tissue targets become outdated after interventional deformation. Our findings achieved adequate 3D localization less than 2 mm, comparable to standard navigation systems using optical or EM solutions.

The integration of an Artis zeego with a *da Vinci Si* represented the state-of-the-art system in currently available clinical CBCT imaging and robotic platforms and was therefore, more extensively analyzed. However, a single-port research *da Vinci* model, the *Sp*, allowed us to explore higher difficulty workspace configurations in additional applications (e.g., gynecology, hepatic). As a result, in gynecology, we experimentally verified that a *da Vinci Sp-zeego* workspace supported a fluoroscopic scan range that was twice that of a *da Vinci Si-zeego* (Table 4.3), and thus could potentially benefit more from improved partial digital tomosynthesis, along with a wider image disparity for X-ray-based 3D localization.

TABLE 4.3 SUMMARY OF WORKSPACE ANALYSIS

C-ARM	DA VINCI	SURGERY	SCAN RANGE [°]	PHANTOM
ZEEGO	SI	TRANSORAL	40	PORCINE
ZEEGO	SI	COCHLEOSTOMY	40	CADAVER
ZEEGO	SI	THORACIC	45	PORCINE, ANTHROPORMOPHIC CHEST
OEC 9600	SI	THORACIC	30	ANTHROPORMOPHIC CHEST
ZEEGO	SI	GYNECOLOGY	50	PORCINE
ZEEGO	SI	HEPATIC	0	PORCINE
ZEEGO	SP	GYNECOLOGY	100	HUMAN (MOCK)
ZEEGO	SP	HEPATIC	0	HUMAN (MOCK)

## 4.5 Recapitulation of Contributions

In Table 4.4 we summarize contributions made towards overcoming technical barriers of specific research problems addressed in Chapter 4.

TABLE 4.4 RESEARCH PROBLEMS (SECTION 1.4) WITH CONTRIBUTIONS TO OVERCOME  
TECHNICAL BARRIERS (SECTION 1.5)

- Problem 2: Preoperative Data and Perioperative/Intraoperative Deformation
  - Feasible perioperative/intraoperative systems workspace



configuration for image acquisitions

- Analysis of free workspace configurations

Note: The author conducted the workspace analysis, including experimental protocols, which were jointly developed with Drs. Jonathan M. Sorger and Mahdi Azizian.

- Intraoperative surgical motion

- Integration of C-arm X-ray-based 3D localization

Note: The author applied a 2D-3D registration algorithm as developed by Drs. Mehran Armand and Yoshito Otake for C-arm X-ray-based 3D localization.

- Problem 3: Effective Image Guidance

- Effective systems engineering with evaluation and validation

- Feasible and deployable clinical workflow

## **5 Augmented Reality for Image-Guided Robotic Surgery**

Over the past two decades, along with an increase in minimally invasive approaches to surgery, traditional endoscopic video cameras have also evolved from analog, and monocular to high-definition stereo digital videography. Computer-integrated surgical systems have taken advantage of these trends by processing and enhancing images, effectively making the endoscope a source and recipient of augmented information, displaying anatomical information otherwise invisible to the human eye. In fact, research in the area of mixed paradigm image guidance systems has led to a prolific development of medical augmented reality systems. In addition to video overlay, surgical robotic systems have been augmented in multiple ways, including collaborative constraints<sup>56</sup>, virtual fixtures<sup>212</sup>, and integration of instructional telestrations<sup>213</sup> from expert surgeons. In general, three major interests that motivate augmentation of endoscopic images include: context finding, visualization of hidden structures, and enhancement of images. Endoscopy for minimally-invasive surgery provides a limited

view and thus can be amended with additional context to increase confidence about current and next steps relating back to the surgical plan. Intraoperative imaging has often been applied to expand the scope of sources of information, especially to identify anatomical structures that cannot be distinguished with an optical endoscopic camera. Image processing to illuminate subtle or initially invisible anatomical information can further enhance the existing videography.

Our proposed solution augments the clinician's view of the physical workspace by overlaying subsurface critical anatomical information derived from the preoperative surgical plan. Furthermore, to improve stereoscopic depth perception, the augmented display included an orthogonal view of the virtual scene and dynamic graphical changes relating tool end effector feedback. Thus, for our proposed image-guidance robotic surgery system, we refer to augmented reality not only as video overlay, but also enhanced depth perception and information feedback regarding intraoperative tool positions.

## **5.1 Related Work**

The Milgram reality–virtuality continuum<sup>214</sup> defines a continuous scale between reality, the unmodeled real environment, and virtual reality, the latter being a purely virtual and modeled environment. Augmented reality concerns the extension of the view of a real scene with virtual objects placed at 3D coordinates respective to the real world. It has applications not only in medicine but also in military, automotive, and entertainment media. However, augmented reality in image-guided surgery involves integrating radiologic images of anatomy with the real intraoperative view of the patient's anatomy to create computerized 2D/3D images of real surgical targets for the purpose of

intraoperative decision-making. Augmented reality aggregates complex technology; thus, its usage within a seamless clinical workflow requires careful design of a software architecture that provides consistent services for image fusion, processing, and rendering.

Augmented reality in medicine displays virtual information on real images of the patient, presenting embedded sub-surface critical structures before physically reaching them, without needing to identify their positions with the sense of touch. Furthermore, augmented reality can also be used to teach novice surgeons by superimposing instructional virtual instruments to sketch a specific task on the endoscopic view. Scientists have employed augmented reality for applications in various domains of image-guided interventions. Approaches differ in visualization, from projective and smart displays to head-mounted gear, x-ray, and video overlay<sup>52</sup>, in addition to direct augmentation of the endoscope<sup>47,48</sup>.

Operating microscopes and binoculars for neurosurgery<sup>215,216</sup> have been augmented with semi-transparent mirrors that reflect the virtual image onto the optical path of the real image. Other efforts include augmentations of angiographic images for ophthalmology<sup>217</sup> and operating binoculars for maxillofacial<sup>218</sup> surgery. The first augmented microscope<sup>216</sup>, arguably the first operational medical augmented reality, showed a segmented tumor slice from computed tomography data. Although this inaugural system achieved an impressive re-projection accuracy of 3 mm, updates for changes in the position of the microscope used ultrasonic tracking that required ~20 seconds processing time. Even in current systems, latency continues to be one of the main challenges of augmented reality in medicine.

Ivan Sutherland pioneered head mounted displays by using a tracking system to

register graphics with the real world for an optical see-through head mounted device, named *Virtual Reality*<sup>127</sup>. In 1992, Bajura, *et al.* describes the use of a head mounted display and a tracking system for augmenting surgical context with ultrasound<sup>219</sup>. The *Virtual Mirror*<sup>220,221</sup> is an advanced interactive visualization method that generates additional mirrored views from any desired perspective on the virtual part of an augmented reality scene. Mobile displays, placed adjacent or above the interventional site, can be used to extend standard sources of information. Ranging from simplistic symbolic graphic representations (e.g., gauges, cross-hairs) to detailed projective anatomical overlays<sup>222-224</sup>, these displays have been popularized with the surge of tablets and smart personal consumer products. Consumer optical head mounted displays, such as *Google Glass* (Google, Mountain View, CA), have already been explored for orthopedics<sup>69</sup> and preclinical pediatric surgical applications<sup>128,225</sup>. Disadvantages of projective and mobile displays include the additional cost and the physical workspace required for new equipment within an already crowded surgical arena. Wearable accessories, such as head mounted devices, create further intrusion and are cumbersome to integrate into current workflows.

Direct overlay of anatomical information onto existing visual displays and video sources for surgery are attractive for minimal footprint and natural integration. An early example of augmented reality in image-guided robotic surgery produced a custom stereoscopic laparoscope, tracked in 6 DOF and viewed through a head-mounted device<sup>104</sup>. A digital light projector aids in depth projection and registration purposes. Towards these efforts, stereoscopic augmented reality has been realized in operating microscopes<sup>226</sup> and transoral robotic surgery<sup>227</sup>. The latter, for ergonomic reasons, similar

to the work by Falk *et al.*<sup>228</sup> that used augmented reality in endoscopic coronary bypass grafting, displayed their fused information directly within the visual field of the *daVinci*'s surgeon-side console. In general, either manual methods<sup>227</sup> or external tracking systems<sup>70,75</sup> are employed to update augmented reality in image-guided robotic surgery. Efforts were taken by Hattori *et al.*<sup>110</sup> to combat intrinsic flickering by optical trackers. A moving average method was applied to an infrared camera tracker motion while applying augmented reality from preoperative CT/MR for robotic cholecystectomy. In additional clinical settings, surgeons have successfully integrated a CT-based surgical CAD model of a 3-cm splenic artery aneurysm using a stereoscopic helmet prior to robotic intervention<sup>61</sup>. A four-week postoperative CT showed excellent functional results including a well-perfused and homogenous splenic parenchyma, with the derived 3D model showing the patency of the re-anastomosed splenic artery.

In the majority of prior work presented so far, the use of rigid-body models has proven successful in many applications, particularly when the anatomical structures of interest are bony landmarks, as is the case for select orthopedics or neurosurgery. However, for other surgical specialties where rigid-body models are insufficient, non-rigid (deformable) models are required. Consequently, select surgical specialties use imaging modalities, including ultrasound<sup>229</sup> and cone-beam CT (CBCT)<sup>116,230</sup> to compensate for intraoperative deformation. Studies<sup>231</sup> have demonstrated the value of 3D ultrasound using transesophageal echocardiography probes to autonomously guide intracardiac robots in phantom experiments. For image-guided urologic procedures researchers have also applied intraoperative 3D ultrasound<sup>232</sup> and MRI<sup>233</sup> to track soft tissue deformation. Prostate interventions<sup>234</sup> have used robotic arms to position a needle

guide for transrectal biopsy by tracking a passive fiducial marker using real-time MR. Actuated pneumatic devices, also paired with MR guidance, have been developed for transperineal prostate biopsy and brachytherapy<sup>142</sup>. Mozer *et al.*<sup>235</sup> present robots compatible with fluoroscopy, CT, and MRI, which were used to orient and guide needles for biopsy and ablation in urology. Furthermore, computer vision concepts, such as optical 3D reconstruction using stereoscopic computations<sup>236,237</sup>, have been applied to update and reconstruct deformable models used for intraoperative verification of surgical CAM.

In an effort to compensate for non-rigid beating heart motion in cardiac surgery, the work of Figl *et al.*<sup>62</sup> explored intraoperative registration of ECG and video processing to register a preoperative 4D MR model of the heart. Validation was conducted using a 4D heart phantom and retrospectively recorded patient images. Similarly, in a post-hoc study, Su *et al.*<sup>49</sup> published their work on augmented reality for robotic-assisted partial nephrectomy. Using reconstructed surgical CAD segmented from preoperative three-dimensional CT data, Su *et al.* augmented video recordings from binocular *da Vinci* endoscopes. Initialized with a manual alignment, their intraoperative updates of the non-rigid kidney surface model used iterative closest point method for deformable registration.

Alternatively, intraoperative imaging can also address deformable, non-rigid, registration. Teber *et al.*<sup>103</sup> integrated C-arm imaging to track navigation fiducials, tested using *ex-vivo* pig kidneys, for image-guided laparoscopic partial nephrectomies. While they achieved a relatively low error margin of 0.5 mm, their proposed approach requires the insertion of several (up to five) 1.5 cm needle-shaped fiducials into the organ. The

insertion of barbed needles into the kidney adds to the intraoperative workflow and incurs risk in perforating vascular structures. Non-invasive fiducial placement on organ surfaces or the abdominal/chest wall may be more clinically acceptable.

Integration of augmented reality with intraoperative imaging has also been realized with augmentation of medical image devices or preoperative volumes. Magnetic resonance imaging data have been fused with an external camera view for neurosurgery<sup>238,239</sup> and video-based augmentation for mobile C-arms<sup>184,185</sup>. Stetten *et al.*<sup>240</sup> augment the real time image of an ultrasound transducer onto the target anatomy. Fichtinger *et al.*<sup>241</sup> propose a similar arrangement of a half transparent mirror and a monitor rigidly attached to a CT scanner allowing for in situ visualization of a slice of 2-D CT. Their efforts using similar techniques was extended for MRI<sup>242</sup>, which is challenged with MRI acquisition compatibility requirements.

## 5.2 Stereoscopic Video Augmentation

For TORS, preoperative planning begins by segmenting critical anatomy, including oncologic and key functional structures, as VTK meshes from standard diagnostic CT. In general, segmentation of such volumetric images refers labeling voxels in the image (or some subset of voxels) as a specific anatomical structure. Research on this topic encompasses a broad and active field<sup>243,244</sup> beyond the scope of the work described here. In the current work, we apply relatively simple semi-automatic segmentation techniques initialized with intensity-based thresholds and edge-based snake methods<sup>245</sup> using ITK-snap<sup>244</sup> and 3DSlicer<sup>246</sup> (Brigham & Women's Hospital, Cambridge MA). The segmentation of structures of interest constitutes a *preoperative surgical plan*, which in our experiments included the tongue targets (centroid and



boundary), registration fiducials, oral tongue, and tongue base volume.

We used the visualization toolkit (VTK) to create a virtual scene that rendered the segmented critical structures, 3D cursors, and menu. A virtual VTK scene consists of actors representing objects, renderers that draw the scene, and cameras that maintain the desired perspective views. In our case, we created two VTK cameras that continuously reflected the Cartesian position of the real stereoscopic endoscopes. Using the registration steps described in Section 3.6.2, the visualization component augmented the primary visual field (endoscopy) by blending a projection of the segmented mesh of critical structures, from the view of the tracked endoscopes, with the original raw endoscopic images. This fused information is directly displayed in the stereoscopic viewport of the *da Vinci* surgeon-side console.

Sielhorst *et al.*<sup>247</sup> present an extensive literature review of medical augmented reality. Current issues for augmented reality in surgery can be broadly organized into registration, tracking, and calibration. Registration, defined as the process of relating two or more data sets to each other or a real physical workspace with matching content, is discussed in Section 3.6. Calibration requirements of our system are presented in Section 3.5. Thus, in upcoming sections, we focus more on tracking, a constant challenge as highlighted by prior art, and updates for different features of augmented reality developed for our proposed system.

### **5.2.1 *da Vinci S/Si***

The patient-side cart of the *da Vinci S/Si* supports a rigid stereo endoscope, manipulated by the central robotic arm. Previous groups have used optical tracking to reflect camera motion in order to update video augmentation of robotic endoscopes<sup>109,110</sup>.

A reflective rigid body is attached near the camera, so the required line of sight can be ensured when the endoscope shaft is inside the patient. In contrast, our approach tracks the endoscopic movements implicitly using kinematic information from the *da Vinci* system’s API. Therefore, no additional tracking system is necessary.

The 3D User Interface (3DUI), described in Section 3.3.2, allows direct user manipulation of the virtual surgical CAD/CAM with respect to the endoscope. Our extension of the existing *cisst* software added functionality to maintain a hierarchy of coordinate systems (Figure 5.2) that manages the virtual scene. Virtual objects are stored relative to each branch in the hierarchy and thus can be transformed individually or collectively. Relative transformations between coordinate systems are updated based on user input and registration steps in the workflow. The 3D User Interface coordinate system serves as the root node of this transformation tree. The motion of the *da Vinci* camera is reported by the API as the movement of the endoscope center of motion (ECM) with respect to the remote center of motion (RCM) of the endoscope’s patient-side manipulator. A static rotation of  $180^\circ$  about the y-axis aligns the surgeon’s console to the 3DUI coordinate system. Transformation of the ECM to RCM is updated when the surgeon moves the camera on the SC by computing the forward kinematics of the endoscopic joint positions. Triggered by the camera pedal “press” events and using real-time joint positions, the 3DUI computes the displacement of the endoscopic center of motion and continuously updates  $^{Endoscope}T_{CBCT}$  to maintain the registration of the overlay.

When the *da Vinci* clutch pedal is activated, the surgeon can move the master manipulators while the patient-side manipulators remain stationary. In the 3DUI behavior, we take advantage of this setting by implementing a mode termed “Masters-As-

Mice.” In this mode, the surgeon can access a 3D menu using the master manipulators rendered as 3D cursors. Supported features include fiducial-based registration, visibility toggle, and direct manipulation of the overlaid objects, allowing the surgeon to manipulate and review the virtual scene, as well as refine the overlay registration all within the native 3D viewport of the SC. A single activation of a master manipulator (MTM) (left or right) translates the virtual objects. If both master manipulators are activated, the movement the 3DUI objects follow the rigid transformation of the segment between the master manipulators at the centroid. These motions therefore only displace the virtual objects whose transformations are maintained with respect to the coordinate system of the remote center of motion.

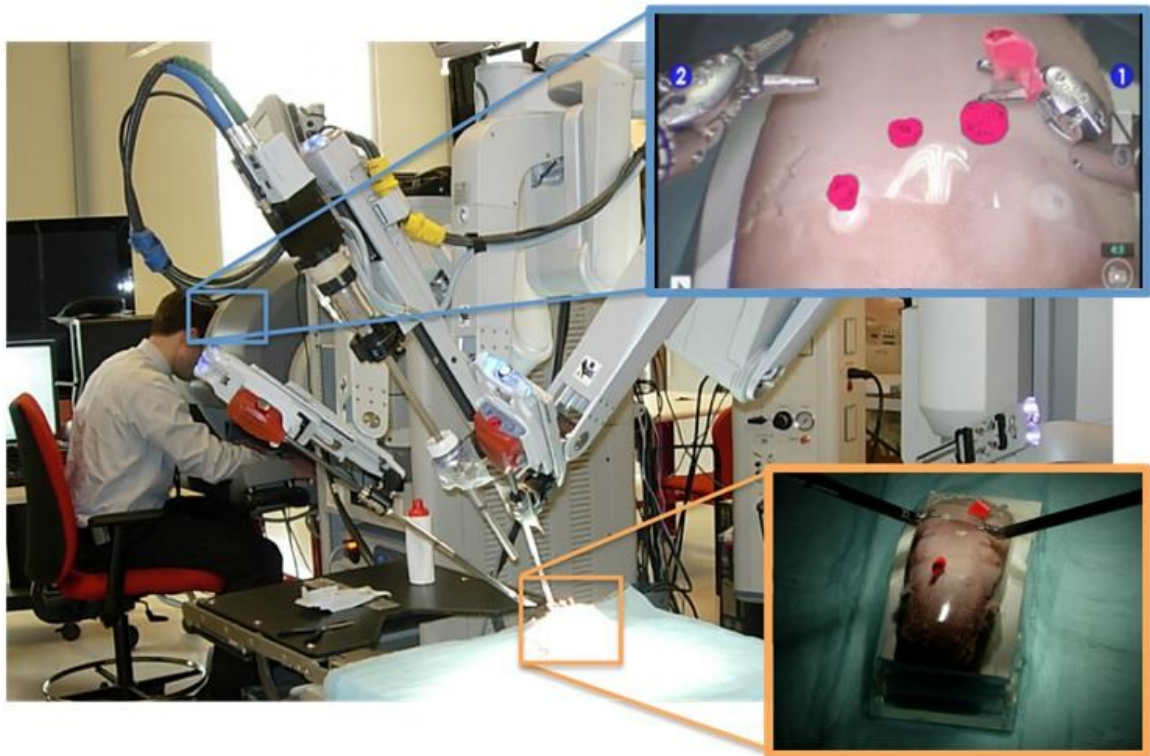


Figure 5.1 Experimental setup using a *da Vinci S* system with stereoscopic video overlay of spherical soft-tissue targets (magenta) segmented from intraoperative CBCT.

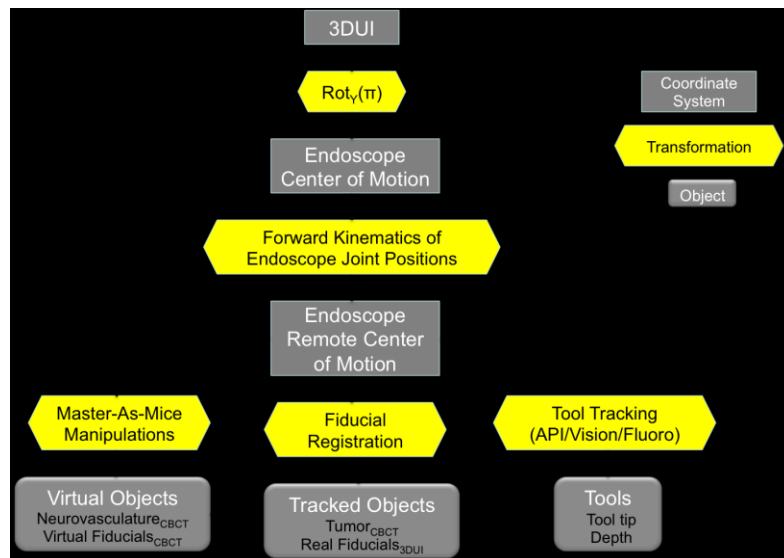


Figure 5.2 A diagram showing the hierarchy of coordinate systems maintained by the 3D user interface.

### 5.2.2 *da Vinci Sp*

Unlike the rigid endoscopes used by *da Vinci S/Si*, the motions of the flexible endoscope of the *da Vinci Sp* (Section 3.3.3.3) cannot be tracked by optical tracking systems. Electromagnetic tracking<sup>248</sup> of the distal tip of a bronchoscope and feature-based tracking<sup>249</sup> have been applied towards navigated bronchoscopy. In fact, for flexible endoscopes, other groups have used kinematic-based, geometric knowledge<sup>250</sup>, and image processing<sup>251</sup>, such as shape from shading or epipolar geometry<sup>252</sup> for updating purposes.

The camera and relative tool motion of the patient-side manipulators for the *da Vinci Sp* have several modes, compared to a single mode of motion for *da Vinci S/Si*. The design of the single port system enforces a shared center of motion located at the trocar incision for all tools and camera; thus, each device must articulate after the insertion point in order to triangulate and create the workspace required.

The predicate *S/Si* EndoWrist instruments of the *da Vinci* are hinged-wrist joints providing 6 degrees of freedom. The EndoWrist of *Sp* instruments has an additional joggle joint in combination with snake-like wrist joints. Joggle joints, are constrained to move together in tandem so that the longitudinal axes are always parallel to each other. Similarly, The EndoWrist *Sp* distal chip camera has the same design as the instruments, including both a wrist joint and joggle joints.

The predicate surgeon-side console (SSC) has a view pedal that activates single control of endoscope position and camera-focusing function. The SSC of the *da Vinci Sp* has a modified view pedal and a new arm pedal.

The three modes of camera motion that the *Sp* system supports are as follows:

- Camera Control Mode (CCM): Using the view pedal, we move the articulated joints on the endoscope while instruments remain immobile.
- Adjustment Mode (AM): Using the view pedal we move the camera through movement of the EGM, in order to re-center the instruments' range of motion.
- Relocate Mode (RM): The arm pedal is used to re-orient the instruments and camera as a group, through the movement of the EGM, pivoting around the single port (i.e., moving the instruments and camera to a different surgical quadrant).

Modes AM and RM, with respect to the optical center of the endoscope for augmented reality purposes, can essentially be considered the same. In these modes, similar to the *S/Si* the da Vinci API reports the transformation of the endoscope with respect to the RCM and can be used directly by our image-guidance system to transform the virtual scene. However, in CCM mode, the API temporarily uses a new coordinate system, situated at the base of the fully retracted camera (patient side manipulator 4, PSM4), and does not switch back until the view pedal is pressed again. Therefore, in order to maintain a consistent transformation of the tip of PSM4 with respect to the RCM,  $^{PSM4}T_{RCM}$ , we compute the transformation chain shown in 5.1 using forward kinematics.

$$^{PSM4}T_{RCM} = (^{RCM}T_{EGM} * ^{EGM}T_{Base} * ^{Base}T_{PSM4})^{-1} \quad (5.1)$$



Figure 5.3 A photograph of the patient side cart of the *da Vinci Sp*, showing the entry guidance manipulator (i.e., the main structural chassis).



Figure 5.4 A photograph of the *da Vinci Sp*, a single-port robotic system with 4 channels within a 25 mm trocar that accommodates a flexible endoscope and 3 articulated instruments.

### 5.3 Tool Tracking

We looked to enhance depth perception through augmented reality by including feedback regarding tool position with respect to critical structures. In this effort, we strived to disambiguate the localization of our tool end effectors along the camera axis (i.e., orthogonal to endoscopic view). Thus, in addition to video overlay, we enhanced our augmented scene with information from calibrated tool tracking for the primary instrument (arm 2 of the patient side cart – i.e., the surgeon’s right handle robotic arm).

The *da Vinci* API reports the tip position of any tool in the coordinate system of the endoscope. However, an unknown intrinsic transformation in the setup joints of the robotic arm offsets the true tool position due to joint encoder uncertainties. To calibrate for this offset, we used two different approaches. First, assuming that a single rigid transformation corrects for this offset, we recorded the pose of our primary tool in 5-10 positions and computed a rigid transformation between the reported position from the API and the observed video coordinates. However, cumulative joint errors in the setup joints may be nonlinear and a one-time rigid correction does not optimally constrain all possible sources of error.

This first approach for setup joint correction was used to track the tip of a map pin that was held by a *da Vinci* large needle driver for the experiments detailed in Section 5.5. To estimate the tip of a pin held by the needle driver during experimentation, we assumed a simple translation (i.e., length of the pin, as measured preoperatively) from the center of the end effectors of the tracked needle driver. These steps allowed tracking and enhanced information using an avatar following the tip of the tracked pin within the endoscopic view and overlaid objects.



A second method used a proprietary computer vision algorithm developed by Intuitive Surgical to provide a continuous correction for SUJ joint errors. This second technique was used to improve tumor resection with margin guidance used in experiments described in Section 6.1.2. The wrist of the primary instrument, a 5-mm monopolar cautery, was covered with a custom-printed black and white pattern (Figure 5.5). Localized with respect to the nearest end effector joint, the patterned marker was tracked with the *da Vinci* endoscopes, using concepts from computer vision. Features of the marker, as seen in both the left and right camera, was used to resolve the parameters for the epipolar geometry of the scene in order localized the 3D transformation of the detected patch. This transformation, along with forward kinematics for the joints distal to marker, continuously resolved the pose of the tool end effector, which was used to update an avatar (Figure 5.5, gray sphere) of the tool tip in the video augmentation in real-time.

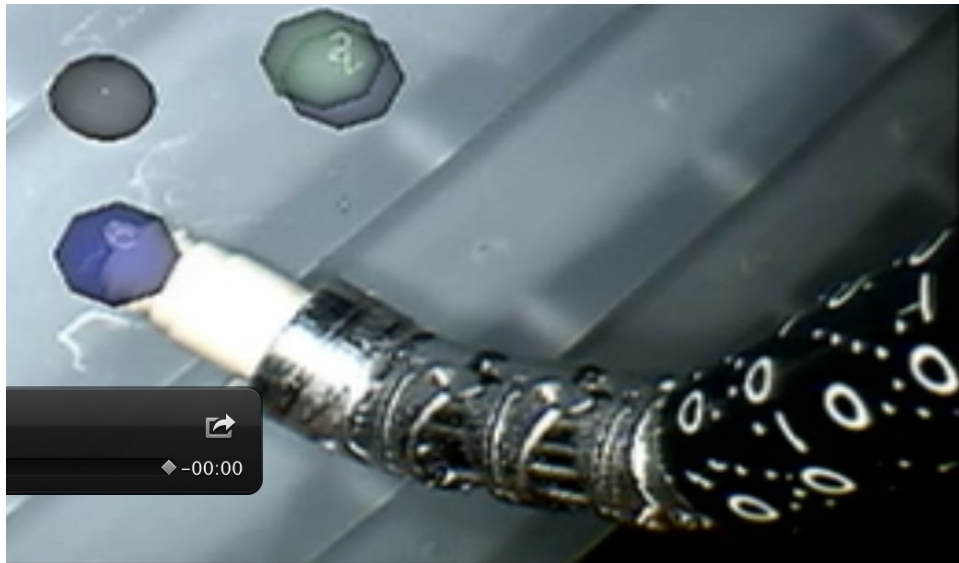


Figure 5.5 Photograph of the feature-based markers attached to a 5 mm monopolar cautery instrument during vision-based tool tracking.

## 5.4 Enhanced Depth Perception

The average optical baseline between a person's eyes measures about six cm. Our visual system therefore takes two slightly different images from different perspectives to project onto the two retinas and be received by different receptors. The corresponding receptors from the pair of images process visual disparity as cues enabling perception of three dimensionality and depth. This depth cue is referred to as stereopsis or binocular disparity.

One of the major issues in augmented reality is correct depth perception in the presence of both virtual and real objects. Incorrect stereopsis has been a topic of discussion since the 1990s when researchers noted natural spatial errors affecting virtual reality as systems portrayed 3D space using a 2D display<sup>253</sup>. When merging real and virtual images, the relative position in depth may not be perceived correctly even though all alignments are accurate. Thus, the usefulness of an augmented reality image-guidance system is a function of both registration accuracy and techniques in depth perception. In fact, multiple experiments in orthopedics<sup>254</sup>, including implant screw placement<sup>131</sup> have shown the benefit from 3D user interaction, as surgeons were able to perform drilling experiments faster with *in situ* visualization compared to a navigation system with a classic 2D display.

Distance from the observer to objects of a scene, as well as distances among objects, is estimated intuitively from learned depth cues. Cues include texture gradients and lighting that affect the human visual system in the perceived properties of objects such as shape, orientation, and dimension. Effects like shadow and reflection on surfaces enrich depth information, while interaction with objects of a scene adds information from

proprioception. This information can be provided as haptic as well as visual feedback. In visual augmentation, the ubiquitous parallax effect occurs when a disparity exists in the apparent position of the object viewed along the two different lines of sight. Nearby objects have a larger parallax and observations of the differences can be used to determine distances. However, if poor calibration or registration amplifies parallax incorrectly, the empirical user observation is that the fused virtual anatomy appears to be detached and floating in front of the real scene. To counter such effects, Bichlmeier *et al.*<sup>255</sup> adjust the transparency according to the position and line of sight of the observer, the shape of the patient's skin and the location of the instrument created a significantly improved fusion of virtual objects to a realistic viewpoint in the scene<sup>255</sup>. Thus, regarding future work for intricate surgical tasks, our hybrid paradigm augmented reality system should investigate improving rendering effects to appropriately combine *in-situ* stereo projective augmented reality and explicit depth perception in order to optimize 3D visualization.

#### **5.4.1 Dynamic Augmentation**

In addition to the projective overlay of the surgical CAD/CAM, we communicated contextual changes corresponding to depth by dynamically manipulating the color of the surface meshes of the critical clinical structures. For target localization in embedded tissue (Section 5.5), we tracked the tip of a pin placed by a patient-side manipulator using a large needle driver. To inject more explicit stereoscopic information, we rendered a transparent sphere at the tracked needle tips that changed from green, to yellow, to red (Sphere, Figure 5.6) when the estimated distance from the pin tip to the closest target was within 4 mm, 2 mm, and 1 mm, respectively. In addition, a quantitative marker (depth

gauge), which followed the tool calipers, displayed the distance to the target (Figure 5.6, White numeric label on end effector). Not only does this give subsurface information but it also resolves ambiguity from perception by quantifying absolute distances.

For the *in vivo* tumor resection experiments described in Section 6.1.2, visual enhancement included an overlay of an ideal margin boundary (i.e., spherical volume with a 10 mm radius, concentric with the tumor). This is facilitated by a change of the sphere's color indicating a breach in the distance between the margin and the tumor. A default blue hue changed to green (note that the large sphere in Figure 6.4 is green since tool tip is in proximity) when the tool tip of the primary instrument was determined to be within +2 mm outside of the margin, then yellow and red when the tip moved within the margin by -2 mm and -4 mm, respectively. In addition to these chromatic cues, a numeric label on the wrist of the instrument was also displayed, thereby showing a real-time update of the relative distance of tool to the ideal margin boundary.

#### **5.4.2 Virtual Perspective**

In addition to recommendations from prior art, we also empirically determined that surgical localization using the *da Vinci* system would benefit from the explicit depth perception in augmented reality. For singular point-to-point relationships, such as the target localization task tested with experiments presented in Section 5.5, we used simple chromatic cues and labels by dynamically modifying existing virtual surgical CAD/CAM. However, in more complex clinically-relevant situations, such as tumor/margin resection, a more comprehensive solution was required.

To provide navigational information in binocular disparity, we needed to disambiguate relative distances of objects in the viewpoint orthogonal to the camera

plane. To address this issue, we implement a novel supplemental view of tracked tools within the virtual scene (model meshes of critical information and CBCT slices and volumes). This auxiliary camera perspective, rendered picture-in-picture (Figure 5.6, lower inset left) can be dynamically changed but is observed to be most useful in the lateral, left-to-right sagittal plane, orthogonal to the camera plane. We implemented the picture-in-picture (PIP) display by extending the OpenIGTLink module for 3D Slicer<sup>132</sup> with a bidirectional interface to the image guidance program. Registration and tool transformations were streamed to the Slicer module, which rendered the tools, CBCT data, and tumor models in the PIP display.

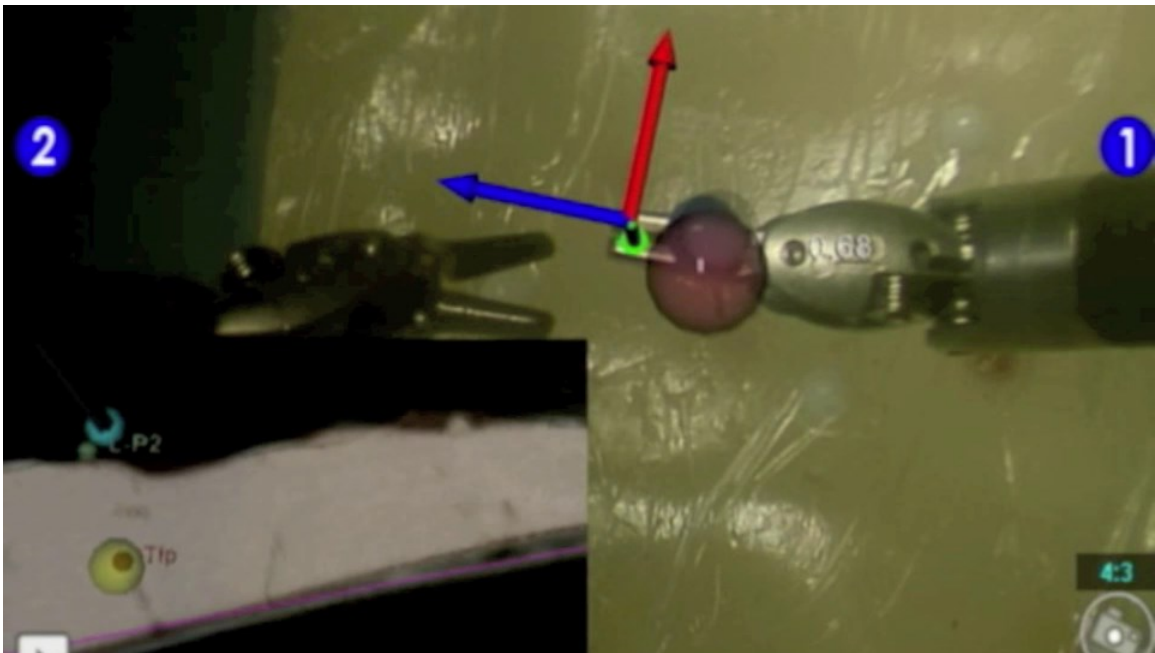


Figure 5.6 Screenshot of the enhanced augmented reality scene with a virtual, orthogonal perspective in the lower left corner. As the tracked needle tip approaches the embedded target, enhancements include the change in color (red) of the sphere, a distance label (white on end effectors), and the virtual perspective (picture-in-picture) showing a red avatar of the tracked point inside the target (yellow).

## **5.5 Preclinical Experiments: Evaluation of Augmented Reality for Transoral Robotic Surgery**

We have proposed the use of intraoperative cone-beam computed tomography (CBCT) to deformably register key anatomical structures delineated from preoperative diagnostic CT or MR imaging. In an effort to evaluate the effectiveness of the previously described methods of augmented reality, we test our system using a simplified target localization task for transoral robotic surgery (TORS).

The following section evaluates our workflow integrating intraoperative cone-beam computed tomography (CBCT) for image-guided TORS through robotic experimentation locating 8-10 embedded targets in five porcine tongues and a cadaveric head phantom. We evaluate several scenarios of image guidance, including experimental variation in augmented reality with video overlay, enhanced depth perception, and tool tracking in comparison to a simulation of current practice.

### **5.5.1 TORS Phantoms**

#### **Model Porcine (Model P) Tongue Phantoms**

Experiments were first conducted on five fresh porcine tongues. Initially, three porcine tongues were each embedded with eight frozen peas simulating soft tissue targets. In the remaining two tongues the soft-tissue-simulating spheres were replaced with ten 1.6 mm diameter Teflon spheres, which resulted in reduced collateral tissue trauma during target placement, and improved accuracy of target analysis. Between eight and ten 3.2 mm diameter nylon spheres were affixed to the tongue surface to serve as registration fiducials. The number of utilizable embedded targets varied due to random

placement and different tongue specimen sizes, which also affected the number of required surface fiducials. The tongue specimens were placed on one of two interchangeable foam templates retained on a custom fabricated frame. The first template maintained the tongue in a flat preoperative (PO) position similar to that of a patient lying supine, while the intra-operative (IO) tongue template placed the tongue in an extended curved position, simulating that of the human tongue retracted during a TORS base of tongue resection. The tongue phantom ( Figure 5.7), including the template and frame, were imaged in positions PO and IO by a C-arm CBCT system with a (15x15x15) cm<sup>3</sup> field of view as detailed below.

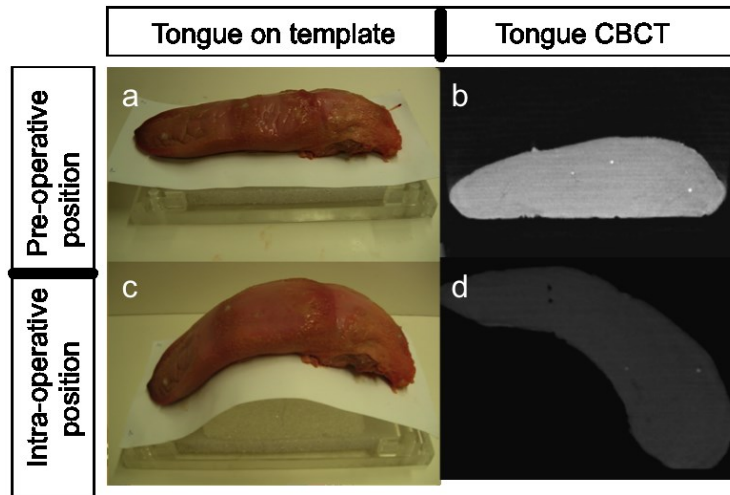


Figure 5.7 Photographs and CBCT images of a porcine tongue positioned by (a,b) a flat, preoperative template and (c,d) a curved, extended intraoperative position. The bright punctuated lesions in the scans represent the Teflon targets.

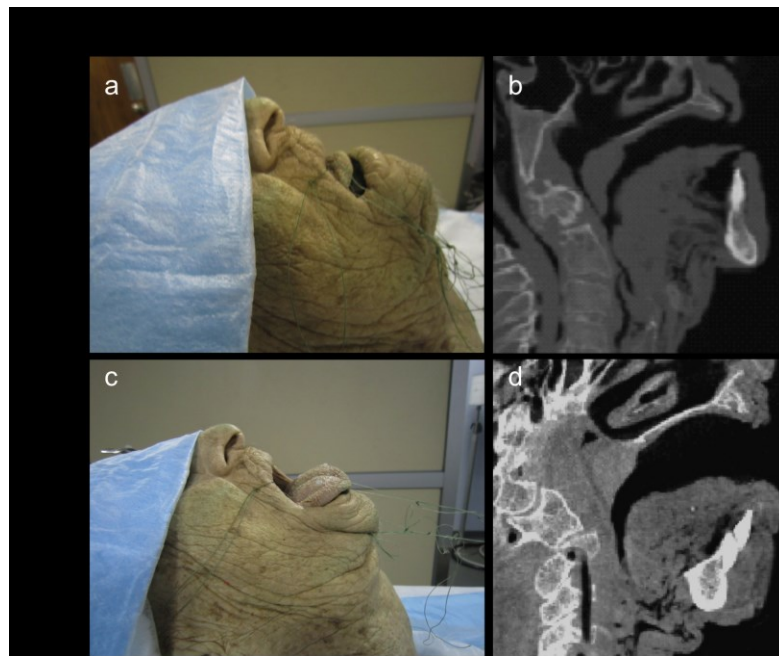


Figure 5.8 Cadaver head positioned in (a) a preoperative pose and (b) imaged in CT. The same cadaver was positioned in (c) an intraoperative pose [tongue sutured and extended] (d) and imaged using C-arm CBCT.



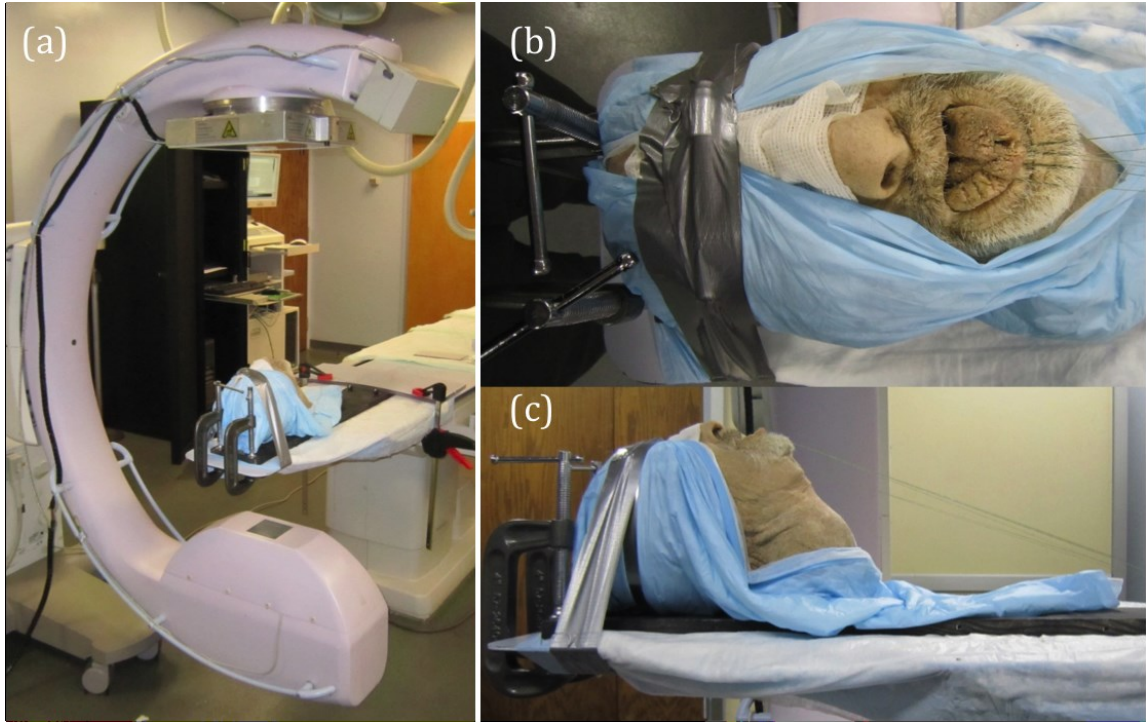


Figure 5.9 Cadaver setup emulating the intraoperative setup for CBCT-guided TORS. (a) Prototype mobile C-arm for intraoperative CBCT. Cadaver mounted in a CT-compatible frame using nylon strings to retract the jaw and tongue as shown in (b) and (c).

### **Model Cadaver (Model C) Head Phantom**

A fresh adult cadaver head from the Maryland State Anatomy Board was lightly preserved in a phenol-glycerin solution to maintain joint and tissue flexibility, allowing for a range of motion in the neck, mandible, and tongue. The cadaver was mounted on a CT-compatible board incorporating nylon strings to open the mouth and retract the tongue in a realistic intraoperative pose (Figure 5.9 b, c) without the large metal retractor and tongue blade that are typical of a clinical TORS procedure. In order to avoid scatter artifacts in CBCT these metallic objects were not used. Modifying such devices to a CT-compatible form (e.g., Al, Ti, or carbon fiber) and applying various artifact reduction algorithms are beyond the scope of the work presented. For analysis of TRE, six 1.5 mm diameter Teflon spheres were glued to the surface of the tongue and eight 1.5 mm Teflon spheres were implanted within the tongue.

In the preoperative pose, simulating that of a patient in a CT scanner, the cadaver's mouth is closed with tongue in repose. In the intraoperative position, replicating that of a patient positioned for base of tongue TORS, the neck is extended, mouth open, and the tongue pulled anteriorly with sutures along a custom radiolucent frame. These positions are illustrated in Figure 5.8. Six 3.2 mm diameter nylon sphere fiducials are glued to the surface of the tongue while ten 1.6 mm diameter Teflon spheres are implanted within the tongue to serve as targets. Clinical mouth and tongue retractors, such as the Feyh Kastenbauer (FK) retractor (Gyrus ACMI/Explorent GmbH, Tuttlingen Germany), were not used as these stainless steel instruments would cause metal artifacts in CBCT. Alternatively, these challenges could be addressed in future work by

potentially using custom radiolucent instruments or advanced reconstruction algorithms with metal artifact reduction<sup>256</sup>.

### **5.5.2 Robotic Experimental Protocol**

A fellowship-trained head and neck surgeon experienced in TORS was asked to use the research *da Vinci S* console with variations of the image guidance system described above. The goal of each experiment was to place 9.5 mm pins using a right-handed needle driver as close as possible to the center of each target embedded in the Model P and Model C phantoms. Experiments on both models included a control, viewing preoperative image data independently from robotic visualization, as well as one or more image-guided scenarios. Experiments conducted on Model P (P1-P5), tested variations of image guidance such as the influence of deformable registration and depth information. In contrast, experiments on Model C (C1-C2) used a more realistic workspace with a cadaveric head to compare simulated current workflow and the full proposed image guidance system. Variations in the experiments are summarized in Table 5.1.

#### ***Model Porcine Experiment***

Experiment P1 simulated current clinical practice, in which preoperative images are viewed separately from the robotic system. The preoperative image data were available offline in standard 3D triplanar views on a laptop next to the surgeon's console. The surgeon had free access to the PO images prior to placement of each needle in its designated target. In experiment P2, the IO image data that accurately represents the deformed tongue in the intra-operative set-up replaced the PO data. The inclusion of scenario P2 allowed direct comparison to the case where IO imaging is available but is

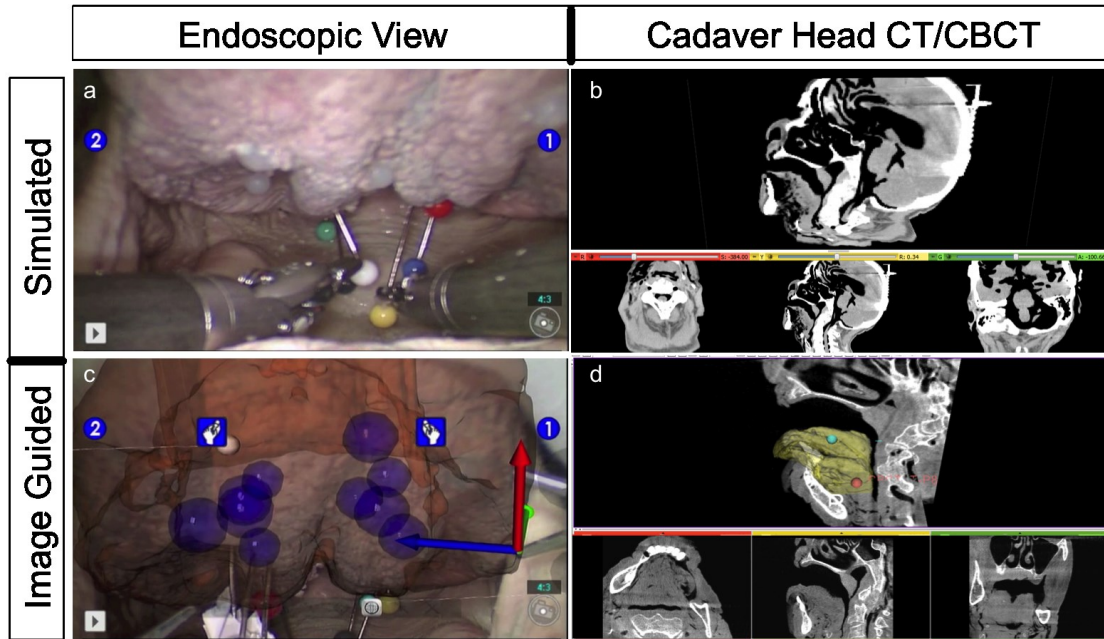
not integrated (i.e., viewed separately from the stereoscopic endoscope). Scenario P3 deformably registered intraoperative models onto the stereoscopic view. This allowed the surgeon to view a 3D overlay of the target images on top of the standard endoscopic view, allowing for the evaluation of the influence of image guidance with overlay and the integration of deformable registration. Scenario P4, or planning directly from CBCT, is an unrealistic situation as preoperative CTA and MRI are expected to better delineate soft tissue oropharyngeal structures. The inclusion of this variation serves as a basis of comparison to an ideal gold-standard situation where image-based registration errors do not contribute to the overall inaccuracy. However, arguments for further exploration comparing CBCTA to preoperative image data can be interesting next steps, as CBCTA would capture patient deformation from setup by anatomical changes since diagnostic image acquisitions. Planning structures were defined directly in the IO images, rather than deformably registered as in P3. Experiment P5 extended the image guidance scenario from P4 with an additional depth gauge to evaluate the impact of explicit stereo information.

### ***Model Cadaver Experiment***

Similar to P1, Experiment C1 simulated current clinical practice, in which standard 3D triplanar views of preoperative CT are available offline for pin placement. Image guidance for Experiment C2 tested the most realistic phantom with the full proposed workflow, using video augmentation with additional depth information of deformably registered preoperative CT data.

Following each of the above experimental protocols and after all pins were placed, the Model P and Model C specimens were imaged with the CBCT C-arm. The

pin tips, pin axis, and targets were manually segmented in CBCT using ITK-Snap and 3D Slicer. Target Localization Error (TLE) was measured as the distance between the pin tip and target. We decompose TLE into four types: Edge, Center, Projection, and Depth, as defined in Figure 5.11. The  $TLE_{Edge}$  is the distance between the needle tip and the closest edge of the target (needles placed on or inside the target were assigned a  $TLE_{Edge}$  value of 0 mm). The  $TLE_{Center}$  best captures the given task of needle placement at the centroid, though  $TLE_{Projection}$ , or reprojection distance, has been used in previous work to evaluate accuracy of video augmentation<sup>257</sup>. If the needle tip, the center of the target, and the projection of the target onto the needle axis are labeled  $P$ ,  $Q$ ,  $R$ , respectively, then  $TLE_{Center}$  is the line segment  $PQ$ . We deconstruct  $PQ$  into segments along the needle,  $PR$ ,



and orthogonal to the needle,  $RQ$ , as  $TLE_{Depth}$  and  $TLE_{Projection}$ , respectively.

Figure 5.10 Simulated current practice in which the surgeon has access to only the (a) raw endoscopic image or (b) preoperative CT to help guide the surgical navigation. Our

proposed image guidance workflow showing (c) a 3D image overlay of the targets on the tongue derived from the (d) preoperative plan registered to intraoperative CBCT.

### 5.5.3 Results

#### Model Porcine Phantom Experiment Results

Measurements of TLE from the Model P experiments with the TORS robot are summarized in Table 5.2 with box plots in Figure 5.12. The mean  $TLE_{Edge}$  [mm], improving progressively in each image guidance scenario, was evaluated at  $4.8 \pm 4.0$ ,  $3.9 \pm 2.9$ ,  $3.2 \pm 3.6$ ,  $2.2 \pm 1.9$ ,  $1.3 \pm 1.2$  for P1-P5, respectively. Similarly, the mean  $TLE_{Center}$  [mm] improved from  $9.8 \pm 4.0$ ,  $8.9 \pm 2.9$ ,  $7.1 \pm 2.8$ ,  $6.7 \pm 2.8$ ,  $5.3 \pm 1.3$  for P1-P5, respectively while achieving a p-value = 0.0151 between P1 and P5. Comparable results were measured for the deformable workflow (P3) and the direct intraoperative overlay (P4), and the lack of statistical significance for all TLE measures between these two cases (e.g., p-value = 0.7036 for  $TLE_{Center}$ ) suggested that the proposed deformable image registration system approaches the ideal scenario of planning directly in intraoperative CBCT.

$TLE_{Center}$  was further deconstructed into its projection and depth components. Experiments on porcine phantoms show  $TLE_{Projection}$  improving by  $\sim 3$  mm between no overlay [ $(6.4 \pm 3.3)$  mm for P1, and  $(6.12 \pm 1.7)$  mm for P2] and scenarios with overlay [ $(3.2 \pm 1.6)$  mm for P3, and  $(3.2 \pm 1.6)$  mm for P4]. However,  $TLE_{Depth}$  showed no improvement comparing the same two scenarios with and without overlay (Figure 5.12 box plot of  $TLE_{Edge}$  in the upper right). This result points to intraoperative imaging as the source for the improvement in depth error, as opposed to an effect from video augmentation, whose influence clearly plateaus. Stereoscopic overlay localizes targets

within the camera image plane well, but it does not provide clear depth localization along the camera axis (i.e., orthogonal to the camera plane). With additional augmentation of explicit depth information (i.e., experiment P5),  $TLE_{Depth}$  is reduced to  $(3.5 \pm 2.0)$  mm. This demonstrated the usefulness of enhancing stereo perception with information along the camera axis. Using stereoscopic video augmentation and a depth gauge, the surgeon was able to place needles with a hit ratio of 25% in experiment P5, compared to 0% for P1 and P2. Improvements between P1 and P2 indicate the positive influence of intraoperative imaging, and the larger improvement between P2 and P3 reinforces the value of visualizing guidance information directly in the surgeon's natural endoscopic window.

### **Model Cadaver Experiment Results**

Similar to the results detailed above for the porcine phantom, the Model C (cadaver head) data using the proposed image guidance system (Experiment C2) showed improved TLE in all four categories, as shown in the box plots of Figure 5.13 and by comparisons of Experiments C1 and C2 in Table 5.2.  $TLE_{Center}$  improved from  $(11.2 \pm 5.0)$  mm for Experiment C1 to  $(5.8 \pm 2.5)$  mm for scenario C2 ( $p = 0.0189$ ). This demonstrates a statistically significant improvement between simulated current practice (C1 – unregistered preoperative imaging) and the proposed image guidance process (C2 – deformable registration of planning data via intraoperative CBCT and overlay in endoscopic video). Comparing experiment P1 (porcine model) to C1 (cadaver), it is clear that the baseline level of difficulty was higher for the latter, which is likely due to increased difficulty in visualization and maneuverability from a realistic robotic setup for TORS in the cadaver oral cavity. The endoscopic camera (0 degree endoscope) was

obliquely positioned in relation to the tongue surface when passed through the mouth. When viewed through the console, this present crowded targets with more overlap and created a more challenging space for localization both visually and spatially. However, the comparable results from experiment C2 (compared to P5) show that the proposed image guidance system overcomes the additional challenges presented by the realistic setup in Model C.

TABLE 5.1 EXPERIMENTS FOR PHANTOM TARGET LOCALIZATION

Phantom	Image Data	Overlay	Depth	Image Guidance	
Model Porcine	P1	Preop	No	No	Preop image is displayed in 2D, offline. This simulates current practice where preop data is viewed separately from the surgeon's console.
	P2	Intraop	No	No	Intraop CBCT is displayed in 2D, offline. This simulates availability of imaging from intraop C-arm without integration with stereo endoscopy.
	P3	Preop (Registered)	Yes	No	Preop segmented plan is deformly registered to stereo endoscopy. This represented the baseline proposed image guidance workflow with video augmentation only.
	P4	Intraop	Yes	No	Intraop segmented plan is overlayed onto stereo endoscopy. This simulates the ideal scenario where TORS critical structures are sufficiently visible and can be segemented in intraoperative CBCT, thus excluding the need for deformable registration between preop and intraop imaging*.
	P5	Intraop	Yes	Yes	<b>Intraop segmented plan and explicit depth gauge is overlayed onto stereo endoscopy. This extends image guidance used in P4 with augmented depth information (i.e. tracking the tip of the pin with respect to targets).</b>
Model Cadaveric	C1	Preop	No	No	Preop image is displayed in 2D, offline. This simulates current practice, in a cadaveric oropharyngeal workspace, where preop CT is viewed separately from the surgeon's console.
	C2	Preop (Registered)	Yes	Yes	<b>Preop segmented plan and explicit depth gauge is deformly registered to stereo endoscopy. This tests our entire proposed image guidance system (augmentation + depth) in a cadaveric oropharyngeal workspace.</b>

\* This is an unrealistic situation as preop CTA and MR are expected to better delineate soft tissue oropharyngeal structures that may not be visible in intraop CBCT. However, this serves as a basis of comparison where image-based registration errors do not contribute to overall accuracy.



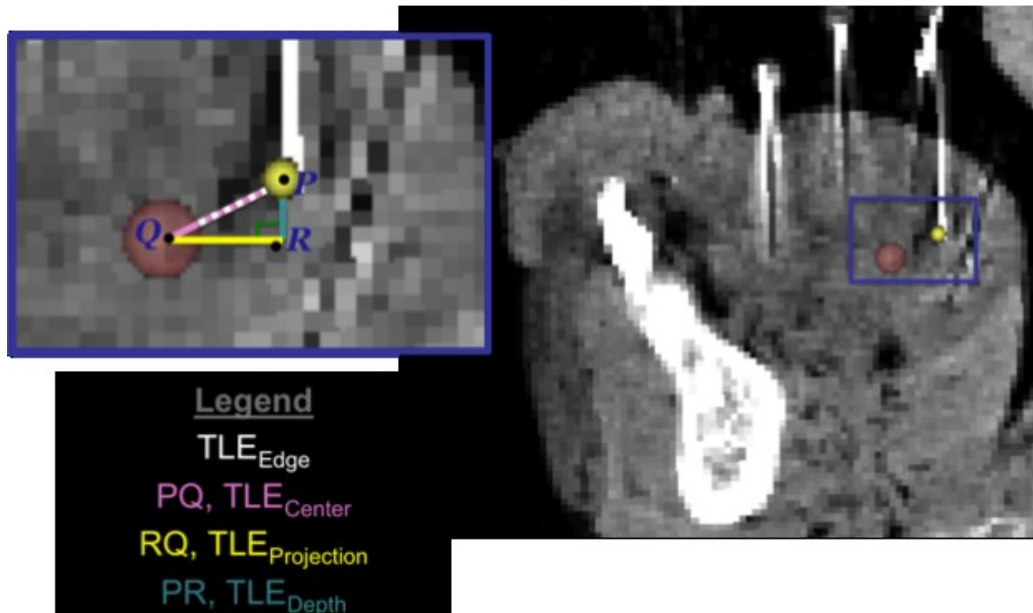


Figure 5.11 Sagittal slice of post experiment CBCT from the Model C phantom (cadaver head). Target Localization Error (TLE) is deconstructed into 4 types: TLE (Edge, Center, Projection, and Depth).

TABLE 5.2 RESULTS FROM PHANTOM TARGET LOCALIZATION

Phantom	$TLE_{Center}$ [mm]			$TLE_{Edge}$ [mm]			$TLE_{Projection}$ [mm]			$TLE_{Depth}$ [mm]			Hit Ratio
	Mean	STDDev	Max	Mean	STDDev	Max	Mean	STDDev	Max	Mean	STDDev	Max	
P1	9.8	4.0	18.5	4.8	4.0	13.5	6.4	3.3	11.3	6.6	4.0	14.6	0/8
<b>P2</b>	<b>8.9</b>	<b>2.9</b>	<b>15.4</b>	<b>3.9</b>	<b>2.9</b>	<b>10.4</b>	<b>6.2</b>	<b>1.7</b>	<b>8.7</b>	<b>5.5</b>	<b>4.1</b>	<b>14.5</b>	<b>0/8</b>
<b>P3</b>	<b>7.1</b>	<b>3.7</b>	<b>14.1</b>	<b>3.2</b>	<b>3.6</b>	<b>10.1</b>	<b>3.2</b>	<b>1.6</b>	<b>5.8</b>	<b>5.9</b>	<b>4.3</b>	<b>12.8</b>	<b>1/8</b>
<b>P4</b>	<b>6.7</b>	<b>2.8</b>	<b>9.8</b>	<b>2.2</b>	<b>1.9</b>	<b>4.7</b>	<b>3.3</b>	<b>2.0</b>	<b>6.0</b>	<b>5.3</b>	<b>3.3</b>	<b>9.3</b>	<b>2/8</b>
<b>P5</b>	<b>5.3</b>	<b>1.3</b>	<b>7.0</b>	<b>1.3</b>	<b>1.2</b>	<b>3.0</b>	<b>3.5</b>	<b>1.1</b>	<b>4.7</b>	<b>3.5</b>	<b>2.0</b>	<b>6.5</b>	<b>2/8</b>
C1	11.2	5.0	19.0	7.4	4.7	15.0	8.1	5.6	16.4	8.1	2.3	10.4	1/10
C2	5.8	2.5	10.1	2.0	2.3	6.1	3.5	1.3	5.5	2.3	1.1	3.2	3/10

TABLE 5.3 STATISTICAL SIGNIFICANCE (P-VALUES) IN THE MEASURED DIFFERENCES IN TLE BETWEEN THE VARIOUS MODES OF OPERATION IN THE PORCINE PHANTOM EXPERIMENTS.

STATISTICALLY SIGNIFICANT RESULTS (P-VALUE < 0.05) ARE IN UNDERLINED BOLD.

**Porcine Experiments (p-value)**

TLE (Center)					
Phantom	P1	P2	P3	P4	P5
P1	-	<b>0.6313</b>	<b>0.0657</b>	<b><u>0.0139</u></b>	<b><u>0.0151</u></b>
P2	-	-	<b>0.1418</b>	<b>0.2569</b>	<b><u>0.0199</u></b>
P3	-	-	-	<b>0.7036</b>	<b>0.2596</b>
P4	-	-	-	-	<b>0.2214</b>
P5	-	-	-	-	-
TLE (Edge)					
Phantom	P1	P2	P3	P4	P5
P1	-	<b>0.6313</b>	<b>0.0831</b>	<b><u>0.0159</u></b>	<b><u>0.017</u></b>
P2	-	-	<b>0.1642</b>	<b>0.2734</b>	<b><u>0.0207</u></b>
P3	-	-	-	<b>0.8183</b>	<b>0.1647</b>
P4	-	-	-	-	<b>0.2109</b>
P5	-	-	-	-	-
TLE (Projection)					
Phantom	P1	P2	P3	P4	P5
P1	-	<b>0.858</b>	<b>0.0588</b>	<b><u>0.0258</u></b>	<b>0.0439</b>
P2	-	-	<b><u>0.0054</u></b>	<b><u>0.0159</u></b>	<b><u>0.0124</u></b>
P3	-	-	-	<b>0.9313</b>	<b>0.6498</b>
P4	-	-	-	-	<b>0.516</b>
P5	-	-	-	-	-
TLE (Depth)					
Phantom	P1	P2	P3	P4	P5
P1	-	<b>0.6179</b>	<b>0.4015</b>	<b>0.6471</b>	<b>0.0887</b>
P2	-	-	<b>0.8948</b>	<b>0.8221</b>	<b>0.2702</b>
P3	-	-	-	<b>0.6957</b>	<b>0.336</b>
P4	-	-	-	-	<b>0.2156</b>
P5	-	-	-	-	-

Table 5.4 Statistical significance (p-values < 0.05) were achieved in measurements of all TLEs in the cadaver specimen between the conventional mode of operation (C1) and the proposed workflow (C2) integrating intraoperative imaging and endoscopic overlay.

### Cadaveric Experiments (p-value)

TLE	Center	Edge	Projection	Depth
Phantom	C2	C2	C2	C2
C1	<u>0.0189</u>	<u>0.0136</u>	<u>0.0406</u>	<u>0.0111</u>

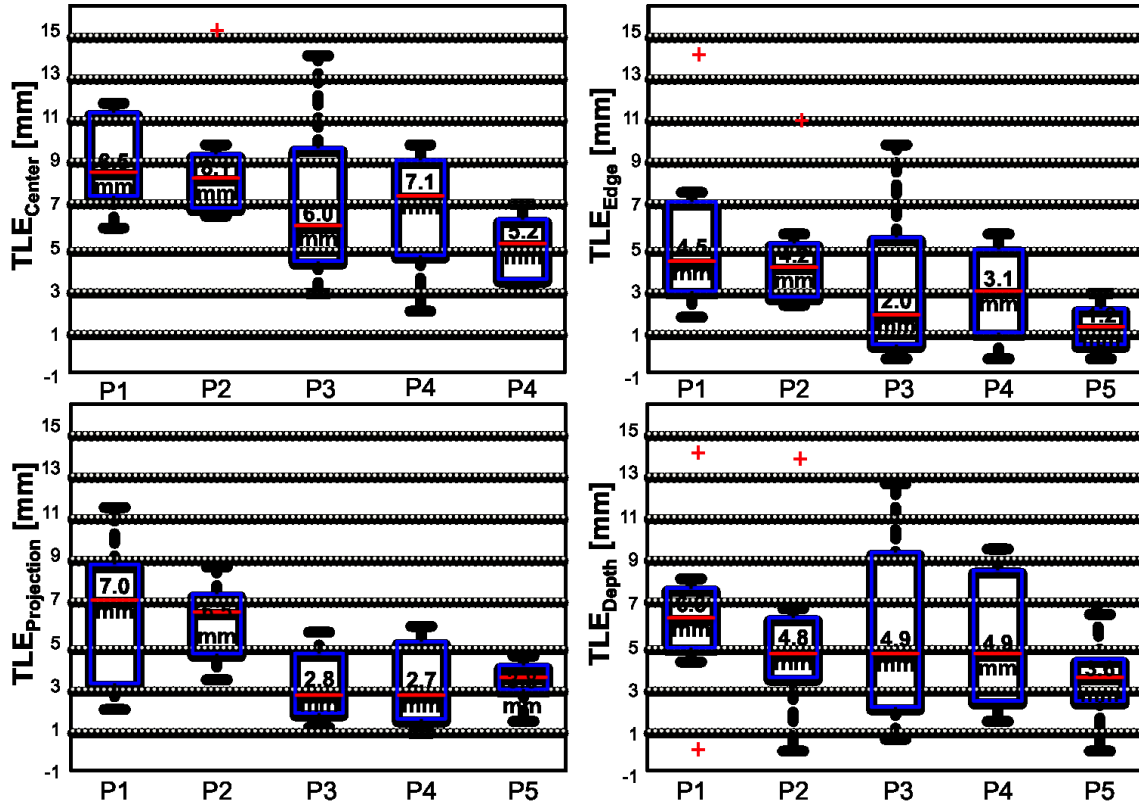


Figure 5.12 Clockwise from upper left, box plots of TLE for Center, Edge, Depth, Projection for porcine experiments.

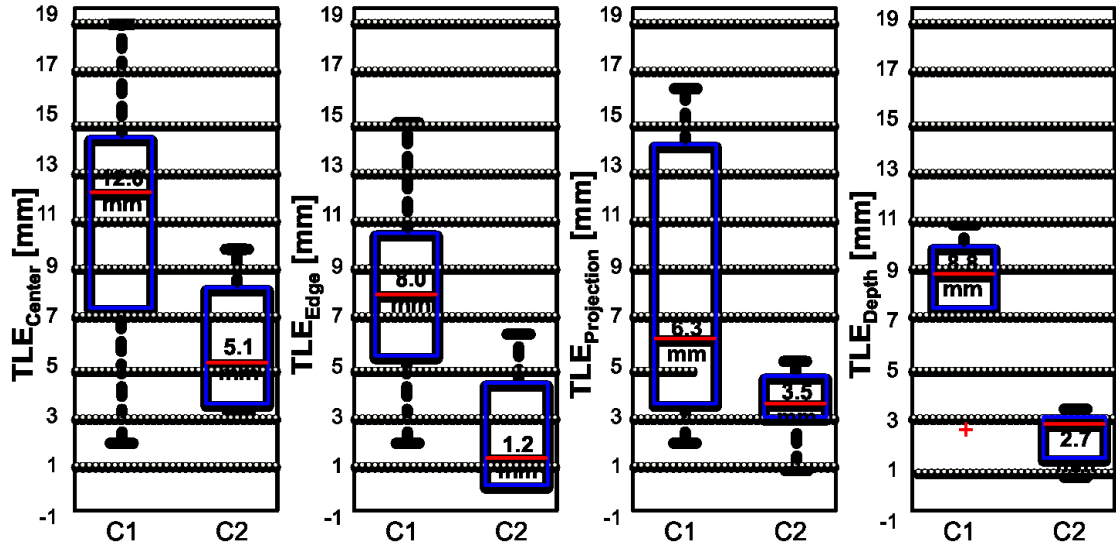


Figure 5.13 From left to right, box plots of TLE for Center, Edge, Projection, Depth for cadaveric experiments.

## 5.6 Chapter Summary and Future Work

Computer-integrated surgery has taken advantage of the evolution of high-definition, digital, monocular, and stereoscopic endoscopy by integrating medical information through augmented reality. Heterogeneous sources of information include modern, rich, diagnostic preoperative data, intraoperative real-time imaging, and robot kinematics (i.e., camera and tool transformations). Thus, applying augmented reality to medicine is mainly motivated by the need to communicate valuable sources of surgical information in the same physical space as the patient. This would improve context sensing and eliminate the current practice of relying on mental correspondence. Along

these efforts, this chapter presented our approach to image-guided robotic surgery through augmented reality, which directs video overlay of critical anatomical information, in addition to enhanced stereoscopic depth information with intraoperative tool tracking within novel orthogonal views of the surgical scene.

Sielhorst *et al.*<sup>247</sup> believed that an optimal *in situ* visualization for medical augmented reality consists of a combination of an augmented window and an head-mounted device. The former allows simultaneous viewing by multiple secondary users, while the latter provides an immersive 3D environment that is visually effective for the primary surgeon. Similarly, we directly augment the endoscope in order to enhance the surgeons' current primary source of visualization as the most "natural window". Due to the design of the vision cart in *da Vinci*-based systems, the enhanced scene is automatically duplicated to auxiliary patient-side monitors. We enhance the projective overlay with additional methods in augmented reality by tracking and integrating local tool end effector positions with respect to the significant virtual objects. We therefore eliminate the ambiguities of the physical spacing, along the axis orthogonal to the camera plane, between the surgeon's tools and anatomical structures of interest. Augmented depth cues, along with the inherent dexterity of the *da Vinci* master manipulators, help to address shortcomings of augmented endoscopes, including improved 3D perception and interaction. Statistically significant results from preclinical phantom-based experiments using our proposed image guidance system show the potential of the various components of our methods in augmented reality to improve target localization for transoral robotic surgery.

A clear limitation of our reported work is the feedback from a single expert user.

For broader distribution, future work can verify these findings and optimize the effectiveness of the visual presentation of our system through a multi-user study. This would identify optimal techniques in rendering, possibly emphasizing contours, lighting, and opacity based the angle of the surface with respect to the endoscope.

## 5.7 Recapitulation of Contributions

In Table 5.5 we summarize contributions made towards overcoming technical barriers of specific research problems addressed in Chapter 5.

TABLE 5.5 RESEARCH PROBLEMS (SECTION 1.4) WITH TECHNICAL BARRIERS RESOLVED  
(SECTION 1.5)

<ul style="list-style-type: none"> <li>• Problem 1: Mental Correspondence <ul style="list-style-type: none"> <li>○ Effective data representation from volumetric CT/MRI/CBCT <ul style="list-style-type: none"> <li>▪ Augmented reality through video overlay of critical anatomies</li> </ul> </li> <li>○ Effective information delivery from tool positions and volumetric CT/MRI/CBCT <ul style="list-style-type: none"> <li>▪ Augmented reality through enhanced stereoscopic depth perception</li> <li>▪ Augmented reality with tool tracking and localization with respect to critical anatomical information</li> </ul> </li> </ul> </li> <li>• Problem 2: Preoperative Data and Perioperative/Intraoperative</li> </ul>
---

## Deformation

- Intraoperative surgical motion

- Computer vision-based tool tracking

Note: The author developed the first method for tool tracking using kinematic information and offset assumption. The second continuous marker-based tracking is a proprietary tool, developed by Intuitive Surgical, Inc. The author contributed necessary software engineering effort to create a deployable module of the tracking software, support for 5 mm tools, and integration into the image guidance system.

- Problem 3: Effective Image Guidance

- Effective systems engineering with evaluation and validation

- Evaluation of effectiveness and accuracy of individual features and methodology of image guidance

## 6 Preclinical Studies: Application

### Validation

Linte *et al.*<sup>108</sup> divide challenges faced by mixed paradigm systems into two broad categories: technical and clinical barriers. From a technical and engineering perspective, each component's inherent limitations and their contributions to the overall navigation accuracy of the system need to be determined. Challenged with appropriate evaluation of robotics and integration of advanced imaging and surgical tracking technology in advanced interventional suites, prior assessments have been conducted on different aspects, including technical studies, the impact on the surgeon's cognitive processes, any changes of surgical strategies and procedures, or impact on patient outcome<sup>258</sup>. Cohen *et al.*<sup>259</sup> apply such an analysis for augmented reality guidance in prostatectomy. Arguably, any system poised for optimal impact on clinical outcome is a well-engineered system with not only accurate registration and tracking, robust visualization, and convincing displays, but also one that fits seamlessly within the standard OR and workflows.

The overall targeting error within an image guidance framework is dependent on the uncertainties associated with each of the components, emphasizing the requirement that a proper validation should estimate the errors at each segment of the entire



pipeline<sup>260</sup> and workflow<sup>261</sup>. Simplified tasks can be used on test materials, or phantoms, manufactured to mimic tissues in mock clinical scenarios. These phantoms enable a large number of experiments to be performed without the complications involved in animal or human studies. Accuracy analysis of our key components can be found as follows: Deformable CT to CBCT Registration (Section 6.1.1), Dual X-ray 3D Localization Accuracy Test (Section 4.3.4); Augmented Reality and Enhanced Depth Perception, Preclinical Experiments (Section 5.5).

However, from a clinical standpoint, the success of an intervention is judged by its therapeutic outcome. Therefore, although the development of such robotic systems requires significant technical innovation and can lead to very real fundamental advances, we needed to show measurable clinical improvements, compared to the standard of care, if they are to be widely accepted and deployed. Using the system in a clinically comparable scenario best assesses potential of the proposed solution. Thus, for each of our motivating clinical applications, we also experimentally evaluate the overall effectiveness of the proposed IGRS system.

## **6.1 Transoral Robotic Surgery**

### **6.1.1 Deformable Image Registration for Cone-Beam CT Guided Transoral Robotic Base-of-Tongue Surgery**

This section presents a summary of experiments, published in 2013 by the Journal of Physics in Medicine and Biology, where Reaungamornrat *et al.*<sup>153†</sup> validate the

---

<sup>†</sup> The deformable registration algorithm was developed by Sureerat Reaungamornrat in collaboration with Wen P. Liu, Drs. Adam S. Wang, Yoshito Otake, Sajendra Nithiananathan, Mr. Ali Uneri, Drs. Sebastian

accuracy of the deformable CT to CBCT registration of the IGRS system for TORS. Significant results include the cadaveric head experiments where base-of-tongue registration accuracy exhibited a median TRE  $\sim 2$  mm, suggesting fairly good geometric accuracy in the central oral tongue and tongue base, the main workspace for TORS.

#### **6.1.1.1 Introduction**

To resolve the large deformation associated with the operative setup in trans-oral base of tongue surgery (i.e., neck flexed, mouth open, and tongue retracted), we propose a deformable registration method that hybridizes a feature-based initialization (using Gaussian mixture (GM) models), followed by a Demons refinement (operating on distance transforms). The combined registration is intensity-invariant and thereby allows registration of preoperative CT and/or MR to intraoperative CBCT.

#### **6.1.1.2 Materials and Methods**

##### **Cadaveric Phantom Head**

A cadaver head specimen, the same as described in Section 5.5.1, was used to verify the accuracy of the deformable registration. Preoperative CT ( $I_0$ ) were acquired (Philips Brilliance CT, Head Protocol, 120 kVp, 277 mAs) and reconstructed at a voxel size of  $(0.7 \times 0.7 \times 1.0)$  mm<sup>3</sup>. Intraoperative CBCT ( $I_1$ ) were acquired using a mobile C-arm prototype (100 kVp, 230 mAs) and reconstructed<sup>262-264</sup> at  $(0.6 \times 0.6 \times 0.6)$  mm<sup>3</sup> voxel size. In the preoperative state, the cadaver is posed with the mouth closed and tongue in a natural pose (Figure 5.9a). Subsequently, the intraoperative pose opens the mouth and

---

Schafer, Jeremy D. Richmon, Jonathan M. Sorger, Jeffrey H. Siewerdsen, and Russell H. Taylor. The author provided technical, and clinical requirements used in the conceptual design of the algorithm.

retracts the tongue in full extension (Figure 5.9b). Radiologic sagittal images of the two poses can be seen in Figure 6.1.

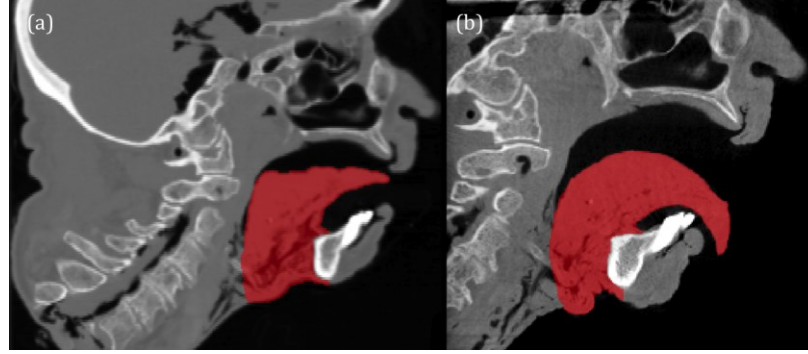


Figure 6.1 Images and segmentation of a cadaver specimen in (a) Preoperative CT ( $I_0$ ) and (b) Corresponding CBCT ( $I_1$ ) in the intraoperative state (mouth open and tongue retracted), with segmentation mask in red.

### Accuracy and Performance

The geometric accuracy of the registration framework is assessed in terms of target registration error (TRE) using two unambiguous anatomical features on the hyoid bone (left and right prominences), in addition to the fourteen Teflon spheres mentioned above. TRE measures the distance between each corresponding target point defined in the fixed image and the moving image after registration:

$$TRE = \|x_1^{target} - p_0^{target}\| \quad (6.1)$$

where  $x_1^{target}$  is a target point defined in the fixed image, and  $p_0^{target}$  is the corresponding target point in the moving image after registration. Thus  $p_0^{target} = x_0^{target}|_{GM_{Rigid}}$ ,  $p_0^{target} = x_0^{target}|_{GM_{NonRigid}}$ , and  $p_0^{target} = x_0^{target}|_{Demos}$  is used to measure TRE following GM rigid, GM nonrigid, Demon, respectively.

A test dataset totaling 25 pairs of CT and CBCT was created by using five preoperative CT images and five intraoperative CBCT images acquired with complete readjustment of the cadaver head between each acquisition. Overall geometric accuracy of registration was measured in terms of TRE.

The total performance runtime of the deformable registration algorithm was measured on a desktop workstation (Dell Precision T7500, Intel Xeon E5405 2x Quad CPU at 2.00 GHz, 12-GB RAM at 800 MHz, Windows 7 Professional 64-bit). For images of size  $(512 \times 512 \times 168)$  and  $(300 \times 320 \times 230)$  voxels for  $I_0$  and  $I_1$ , respectively, the nominal algorithm parameters with a hierarchical pyramid of size [8, 4, 2] used a total runtime of ~5 minutes, including the following distribution: 12 s for point-cloud extraction, 1.3 s for the GM rigid step, 2.2 min for the GM nonrigid step, and 2.6 min for the Demons step. Although the runtime was relatively slow, and there is room for improvement in the implementation and parallelization of the algorithm, it is potentially within logistical requirements of a research clinical workflow.

### **6.1.1.3 Results**

The TRE improved from  $7.8 \pm 7.3$  mm (median 4.4 mm) following GM rigid to  $2.3 \pm 1.1$  mm (median 2.1 mm) following GM nonrigid, and  $1.3 \pm 0.7$  mm (median 1.1 mm) following Demons (Figure 6.2). The semi-opaque overlays of volumes show that the Demons step not only refined the GM registration in deep tissues of the oral tongue, but also in surface matching. The interquartile range in TRE following Demons was 1.06 – 2.21 mm, and the range was 0.09 – 4.51 mm. The fairly broad range and upper bound in TRE can be primarily attributed to target points at the tip of the tongue. Adjustment of registration parameters within the proposed operating range is possible to allow more

deformation and improve overall TRE. This is further analyzed by Reaungamornrat *et al*<sup>153</sup>. Overall the accuracy of hybrid deformable registration is best ( $\sim 1$ - $2$  mm) medial and center in the oral tongue and worst ( $\sim 2.5$ - $3.0$  mm) at the exterior surface tips. Resultant deep caudal aspects of the base-of-tongue registration accuracy test exhibited a median TRE of  $\sim 2$  mm, suggesting fairly good geometric accuracy in the main region of interest for TORS.

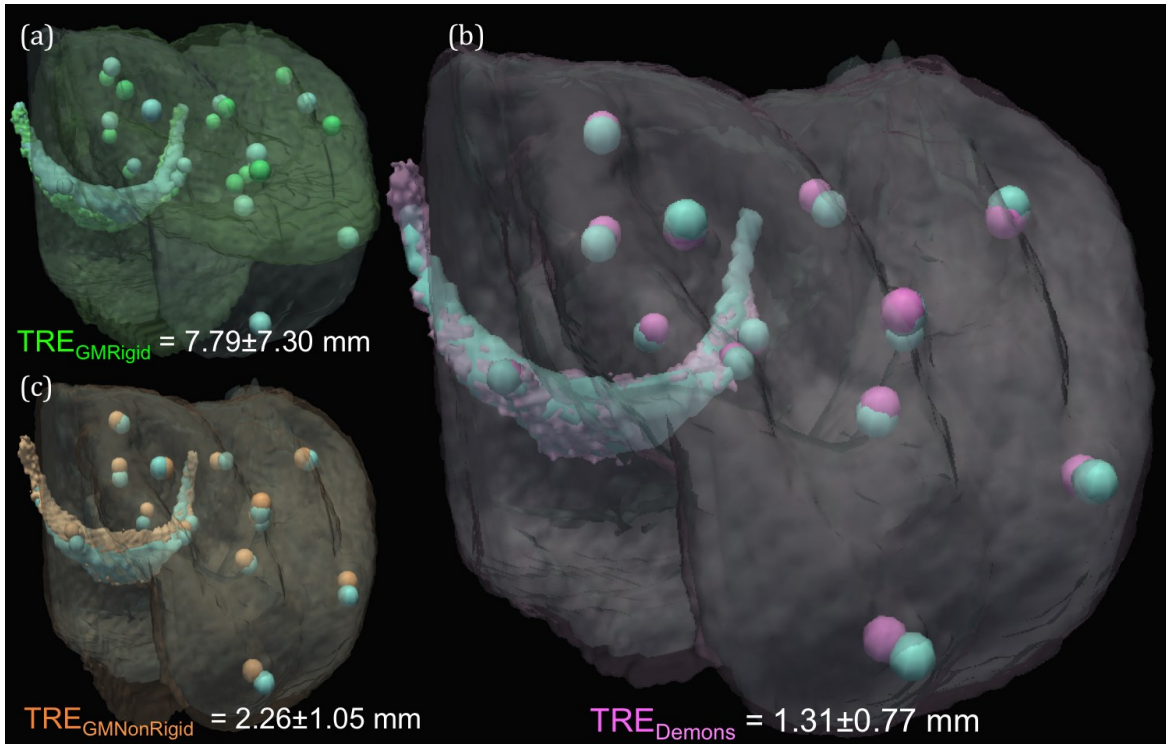


Figure 6.2 The transformed cadaveric tongue and TRE following each step of the registration framework: (a) GM rigid; (c) GM nonrigid; and (b) Demons.

## 6.1.2 Intraoperative Cone-Beam CT Guidance for Transoral Robotic

### Surgery

This section, a modified reproduction of Liu *et al.*<sup>265†</sup>, originally presented at the 2014 Conference for Information Processing in Computer Assisted Intervention, is a summary of preclinical testing of the image-guided robotic surgical system for TORS. This work highlights *in vivo* porcine experiments, where results from resection embedded mock tumors using variations of image guidance are analyzed for achieved margin ratios.

A workflow to overlay critical structures from intraoperative cone-beam computed tomography angiogram (CBCTA) for resection of base of tongue neoplasms is evaluated using *ex vivo* and *in vivo* animal models comparing our image guidance through video augmentation to simulated control and fluoroscopy-based image guidance. Results included visual confirmations of augmented critical anatomy during controlled arterial dissection and successful mock tumor resection. The proposed approach to image guidance also achieved improved resection ratios of mock tumor margins (1.00) when compared to control scenarios (0.0) and alternative methods of image guidance (0.58).

#### 6.1.2.1 Introduction

In contrast to the previously proposed high-level clinical workflow, these experiments propose identifying critical structures directly from intraoperative cone-

---

<sup>†</sup> The development of a system integrating intraoperative CBCTA for guided transoral robotic surgery and validation experiments were completed by the author in collaboration with Drs. Jeremy D. Richmon, Mahdi Azizian, Jonathan M. Sorger, and Russell H. Taylor. Dr. Richmon was the clinical principal investigator and participant for these experiments. Drs. Azizian and Sorger provided technical and academic expertise in system engineering and experimental protocol. Dr. Taylor served as the academic principal investigator for this work.

beam computed tomographic angiography (CBCTA). CBCTA eliminates the need for preoperative planning and deformable registration in addition to the advantage of acquiring an updated volumetric data set of the patient's physiology.

#### **6.1.2.2 Materials and Methods**

##### ***6.1.2.2.1 System Overview and Workflow***

In this experimental workflow, the patient is positioned in a standard intraoperative position, and contrast material is injected to enable visualization of critical oropharyngeal structures while an intraoperative CBCTA image is obtained. Critical data as well as registration fiducials are manually segmented from the CBCTA using ITK-Snap and registered to the stereo video camera of the *da Vinci* robot. Segmentation by intensity-based thresholds (manual initialization) from angiographies can be accomplished on the order of seconds. Alternatively, detailed preoperative planning based on standard diagnostic CT/MRI can be created prior to the operation, which would otherwise contribute to the overall intraoperative time. Guidance through video augmentation (refer to <sup>266</sup> for details of system architecture) is implemented by extending the *SURGICAL ASSISTANT WORKSTATION (SAW)* open-source toolkit<sup>129</sup>, developed at the Engineering Research Center for Computer Integrated Surgery (CISST ERC, Johns Hopkins University, Baltimore, MD). Visual overlay of TORS resection targets (tumor/ margins) and the lingual artery are directly rendered within the endoscopic video to guide the surgeon during base-of-tongue tumor resection. The augmentation follows camera kinematics, provided by the *da Vinci*® application programming interface (API), and intraoperative tracking of custom fiducials. Orthogonal views of tracked tools relative to

the critical data are added to supplement the surgeon's stereo perspective in depth, (i.e., parallel to the camera axis).

#### **6.1.2.2.2 Porcine Models**

##### **Ex vivo (EV) Porcine Tongue Phantoms**

*Ex vivo* excised porcine tongues (Figure 6.3c) were used in simple experimental scenarios. To simulate current standard of practice, as a control scenario, EV models were used in mock tumor resection without integrated image guidance (i.e. CBCT viewed in offline displays). Custom features and settings (i.e., determining color and opacity values for augmented structures and thresholds for tool tracking) for the user interface (UI) was initially tested using *ex vivo* models prior to *in vivo* experiments. Each EV tongue was embedded with a synthetic mock tumor, an 8 mm diameter nitrile sphere (green in Figure 6.3d). Five to eight 3.2 mm diameter nylon spheres (green in Figure 6.3c) were affixed to the tongue surface, which served as registration and landmark fiducials. A CBCT (109 kVp, 290 mA, 0.48x0.48x0.48 mm<sup>3</sup> voxel size) was then acquired with the tongue secured onto a flat foam template.

##### **In vivo (IV) Porcine**

For *in vivo* experiments a live pig is placed supine on an operating table (Figure 6.3d), anesthetized, catheterized, and intubated with a tracheostomy tube. The specimen's jaw was opened with a triangular wooden block wedged between the molars, and the tongue was pulled anteriorly with sutures, configuring the base-of-tongue in an intraoperative position for TORS. An intraoperative angiography was acquired with an injection of 40 ml of iodine (MD-76R) during a volumetric CBCT scan (90 kVp, 290



mA, 0.48x0.48x0.48 mm<sup>3</sup> voxel size). Two mock tumors (Urethane, medium durometer spherical medical balloons, 10 mm in diameter) were placed anterior/ superior to bilateral lingual arteries using a radiopaque FEP I.V. catheter (Abbocath-T 14G x140 mm) in the base of the tongue (Figure 6.3a, b). Balloons were injected with a mixture of 0.5 ml rigid polyurethane foam (FOAM-IT) and 0.25 ml iodine (MD-76R) to retain shape and provide tomographic contrast, respectively. Acrylic paint (0.25 ml) was also added to the filling mixture to provide visual feedback. For S1, eight 3.2 mm diameter nylon spheres (Figure 6.3c) were placed on the tongue surface as registration fiducials, but for S2 these were replaced by a custom resection fiducial (shown in Figure 6.4 as a triangular green lattice with inset white, black and yellow spheres).

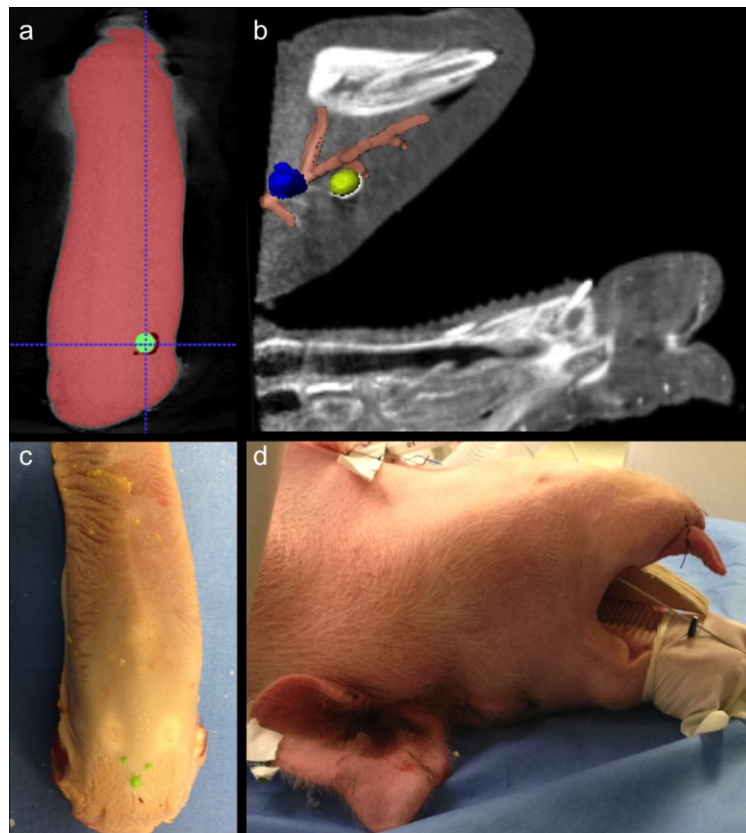


Figure 6.3 (a) Single axial slice from CBCT of an *ex vivo* pig tongue phantom with embedded tumor (green). (b) Single sagittal slice CBCT angiography of an *in vivo* pig phantom with segmented models of the right lingual artery (orange), and two base-of-tongue tumors (right in yellow, left in blue). (c) Photograph of an *ex vivo* pig tongue phantom affixed with green registration fiducials. (d) Photograph of an *in vivo* pig phantom supine and readied for tumor placement.

#### **6.1.2.2.3 Image Guidance**

##### **Video Augmentation and Tool Tracking**

During mock tumor resection, the stereographic projection viewed through the *da Vinci® Si* surgeon side console (SSC) is superimposed with mesh models of critical data segmented from intraoperative CBCTA. Augmentation for EV models included the synthetic tumor and surface fiducials while IV models also included segmented lingual arteries. IV experiments for S2 added a spherical margin (Figure 6.4, green sphere) providing a boundary for the surgeon with the overlay of an ideal margin resection (i.e., spherical volume with a 10 mm radius, concentric with the tumor). This was facilitated by the change of sphere's color indicating a breach in the distance between the margin and the tumor. The default blue hue changed to green when the tool tip of the primary instrument (5 mm monopolar cautery) was determined to be within +2 mm outside of the margin (e.g., the large sphere in Figure 6.4 is green since the tool tip is in proximity), then yellow and red when the tip moved within the margin by -2 mm and -4 mm, respectively. In addition to these chromatic cues, a numeric label on the wrist of the instrument was also displayed, thereby showing a real-time update of the relative distance of tool to the ideal margin boundary.

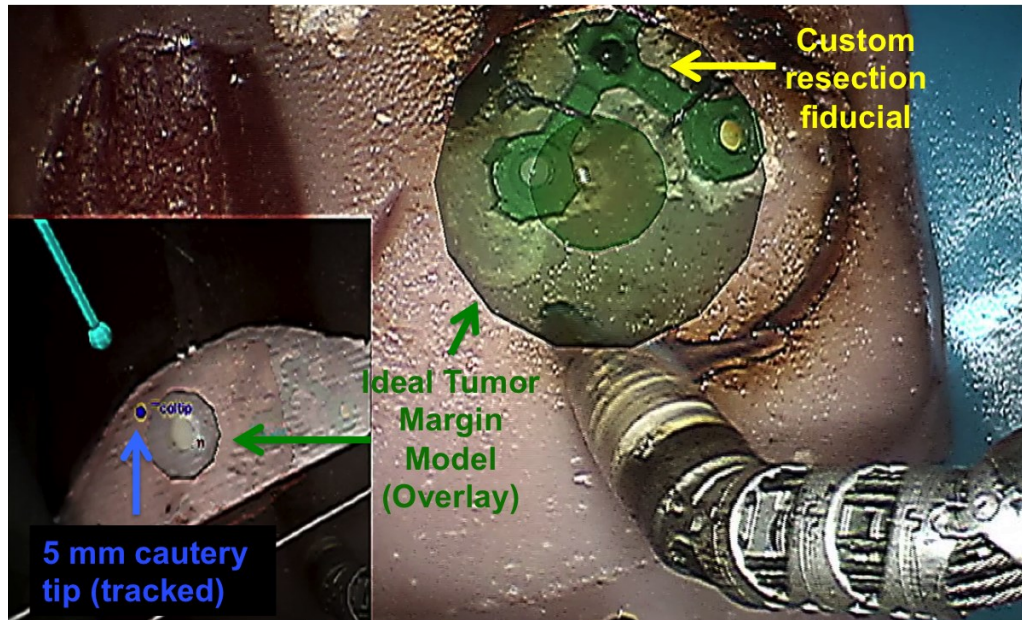


Figure 6.4 Screen capture of an ex vivo phantom experiment using video augmentation of margins (green sphere) and tool tracking in novel views (lower left picture-in-picture) for image guidance.

To provide navigational information in the axis orthogonal to the camera plane (i.e., depth information), we implemented supplemental camera views of tracked tools within the virtual scene (model meshes of critical information and CBCT slices and volumes). This auxiliary camera perspective, rendered picture-in-picture (Figure 6.4, lower inset left), can be dynamically changed but was observed to be most useful in the lateral, left-to-right sagittal plane, orthogonal to the primary axis of the stereoscopic endoscope. We implemented the picture-in-picture (PIP) display by extending the OpenIGTLink module for Slicer 3D<sup>132</sup> (<https://www.slicer.org>) with a bidirectional interface to the image guidance program. Registration and tool transformations were streamed to the Slicer module, which rendered the tools, CBCT data, and tumor models in the PIP display.

For S1, superimposed virtual structures are initially rigidly registered by identifying point-based correspondence with artificial surface spheres (Figure 6.3c), visible in stereo video and segmented from CBCTA. In experiments for S2, as a first step toward intraoperative updates, we continuously track a custom rigid fiducial attached directly above the resection target (Figure 6.4). Assuming a constant spatial relationship within the resected volume (i.e., between the fiducial and targets), we update the overlay of the tumor and margin mesh with the transformation of the tracked custom fiducial. This fiducial was fabricated on a 3D printer and designed as a planar right isosceles triangular lattice with a hypotenuse of 10 mm in length. Each corner of the symmetric triangle was connected by an annulus, a ring with an inner radius of 1.5 mm. The triangular frame (1 mm in width) was painted green, and white, yellow and black 1.6 mm (radius) Teflon spheres are each inserted into corner annuli. Using color thresholds, the green framework of the fiducial was first located as an initial region of interest. Corner annuli of the green frame created circular negatives that were segmented using contours detection, and then matched by their average color to the nylon spheres. Chromatic thresholds, updated on successful fiducial segmentations, were designed to be dynamically adaptive in order to be robust to fiducial color changes due to pollution from cautery. A rigid transformation from point-based tracking of the spheres on the customized fiducial updated the locally rigid transformation of the attached resected volume.

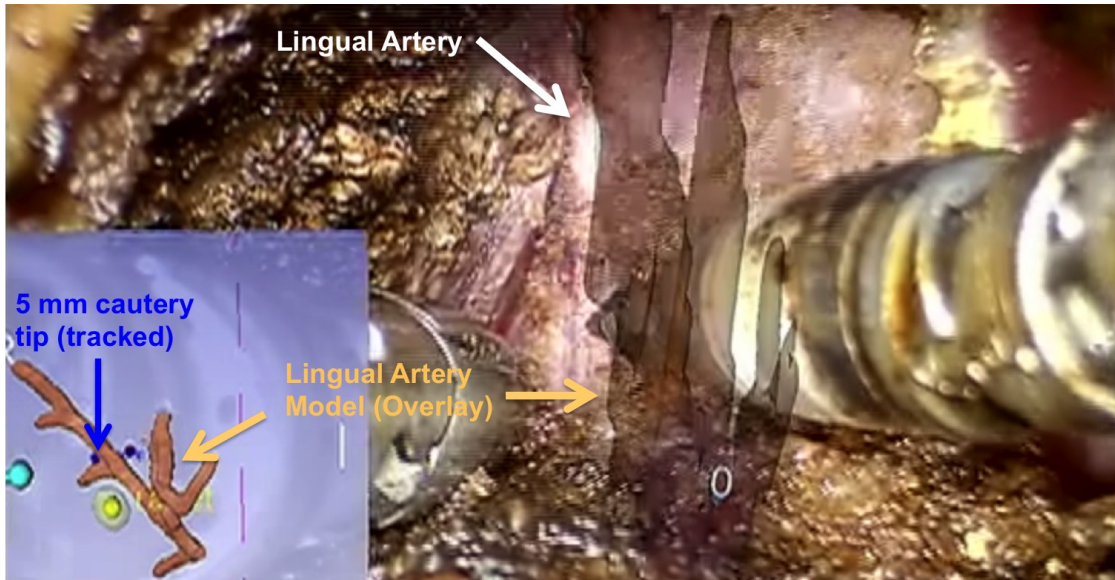


Figure 6.5 Screen capture of lingual dissection during an in vivo porcine lab experiment using video augmentation as image guidance.

Forward kinematics from instrument joint encoders, as provided by the API, has been measured with an error  $\leq 25 \text{ mm}^{267}$ . To correct for this offset, necessary for tool tracking, we tested two methods: 1) Establish the Euclidean transformation corresponding to tool tip locations in stereo video and the API in several tool poses. 2) Derive setup joint corrections through vision-based processing of markers attached to the shaft of the instrument (a proprietary function developed by Intuitive Surgical, Inc.)

### Fluoroscopy Augmentation

For comparison with video-based augmentation, we also tested scenarios with fluoroscopy-based image guidance using the Siemen's syngo workstation. For the IV experiments, after docking the robotic arms to the operating table, the C-arm was placed laterally to capture sagittal X-rays of IV experiments. For intraoperative fluoroscopic-

guided experiments, the surgeon side console was set up in a radiation-shielded workspace with access to manually activated X-ray on request. The live fluoroscopic images and overlay onto the CBCTA of the head of the porcine specimens were rendered in 2D in the bottom left and right corners of the SSC through TilePro<sup>®</sup>.

### 6.1.2.3 Experiments

During two sets of experiments (S1 and S2), a head and neck surgeon<sup>§</sup> resected embedded mock tumors, commensurate with standard surgical practice (i.e., attempting to achieve a 10 mm margin around the tumor while avoiding and/or controlling the lingual artery), from EV and IV phantoms using a research *da Vinci*<sup>®</sup> Si console with variations of the proposed image guidance. Variable scenarios are summarized in Table 6.1. Each set of experiments included: (a) control with EV phantom, (b) video augmentation with EV phantom, (c) Fluoroscopy augmentation with IV on left base-of-tongue and (d) video augmentation with IV on right base-of-tongue.

Two EV specimens, S1a and S2a, were used as controls (i.e., preoperative images were available offline but not integrated to the robotic system) in order to simulate current standards of practice. The clinician was given access to view preoperative CBCTs, with visible tumors and surface landmark fiducials on offline monitors displaying the reconstructed volumes in MPR (Multi-Planar Reconstruction) views. Scenarios S1b and S2b served to gauge user experience and feedback on proposed features of the video augmentation software on simple EV specimens prior to testing on comprehensive IV models. Experiments comparing video to fluoroscopic augmentation

---

<sup>§</sup> Dr. Jeremy D. Richmon, Director, Head and Neck Surgery Robotic Program and Associate Professor of Otolaryngology-Head and Neck Surgery, Johns Hopkins Hospital, Baltimore, MD.

were conducted on IV specimens, which provided a realistic oropharyngeal workspace. Though both S1c and S2c used fluoroscopic augmentation, the S2c setup included the capability to enlarge (4x) regions of interest.

Video augmentation for S1 differed from S2 as follows. For tool tracking, to calibrate for the inherent offset at the remote center of motion S1 (S1b, S1d) used an initial point-based calibration for corrections. A vision-based technique to track artificial markers was employed for S2 (S2b, S2d). In addition, S2 also tested initial implementation to guide margin resection. Augmented overlays of critical data included an ideal margin, updated during intraoperative tracking of a custom resection fiducial.

TABLE 6.1 TORS EXPERIMENTS

Experiment Model		Image Data	Guidance Margin		Image Guidance	
Set 1	S1a	EV	Intraop CBCT	None	No	The control for the first set, simulated current practice with preoperative CBCT shown in offline 2D display, i.e. separate from surgeon's console
	S1b	EV	Intraop CBCT	Video	No	Represented the basic proposed image guidance workflow where stereo video endoscopy was superimposed segmented tumor and fiducials from CBCT along with tracked tools in novel virtual views, picture-in-picture.
	S1c	IV	Intraop CBCTA	Fluoro	No	Superimposed fluoroscopy on CBCTA used for guidance in tumor resection and lingual dissection.
	S1d	IV	Intraop CBCTA	Video	No	Used video augmentation guidance similar to S1b, however with in vivo porcine lab requiring lingual dissection
Set 2	S2a	EV	Intraop CBCT	None	No	The control for the second lab, with the same guidance given as S1a
	S2b	EV	Intraop CBCT	Video	Yes	Included all features implemented for S1b, but extended image guidance for margins with resected specimen tracking using a custom fiducial
	S2c	IV	Intraop CBCTA	Fluoro	No	Fluroscopic guidance, as given in S1c, with extended capabilities to zoom up to 4x on region of interest
	S2d	IV	Intraop CBCTA	Video	Yes	The same guidance and scenario as S1d, but extended image guidance for margins with resected specimen tracking using a custom fiducial

#### 6.1.2.4 Results

For both control scenarios (S1a, S2a showing tumor resection without integrated image guidance on an EV tongue), the resected specimen failed to contain the target mock tumor.

All experiments with integrated image guidance successfully resected the whole tumor. In the live animal lab cases, accuracy of the lingual artery overlays were visually confirmed (Figure 6.5, video augmented overlay of exposed lingual dissection) along with successful arterial dissection and control in both video and fluoroscopic augmentation.

Measurements of the specimen resected from all eight robotic experiments are summarized in Table 6.2. Figure 6.6 shows photographs and corresponding postoperative slices/volumes of the resected tumors (blue in ‘Volumes’ column) and their intersection with an ideal margin (yellow in ‘Volumes’ column). Resection ratios (volume of margin resection/volume of ideal spherical margin) in order from high to low is achieved with S2d (1.00), S2b (0.87), S1d (0.81), S1b (0.71), S2c (0.58) and S1c (0.44), respectively. The challenging environment of a featureless *ex vivo* model, compared to a realistic *in vivo* model, is substantiated with the superior results obtained comparing *in vivo* to *ex vivo* experiments which used the same video-based augmentation for image guidance. Improvements achieved by S2d, in reference to S1d (similarly from S2b to S1b), can be attributed to the addition of margin overlay and intraoperative tracking of the resected volume. S1c and S2c, scenarios that utilized fluoroscopic overlays, had the advantage of precise tool to tumor distances, but were restricted to a single X-ray (2D) plane.

The margin/specimen ratio on S1c and S2d was large secondary to posterior placement of the IV mock tumor and the custom resection fiducial mandated inclusion with the resection specimen. Generally, most IV specimens required longer dissections, resulting in smaller ratios as compared to EV, due to volumes removed for arterial control and workspace limitations of the transoral access.



In addition to resection ratios, two forms of accuracy are of interest here: 1) Projection Distance Error (PDE) – the 2D pixel distance between projected overlay and the true image location of the object; and 2) Tool Tracking Error (TTE) – the 3D position [mm] of the tool tip compared to the tracked virtual overlay. Mean PDE, from point-based manual registration, has been previously established at 2 mm using an anthropomorphic skull phantom<sup>266</sup>. During video-based image guidance for S2 visual estimates of tool tracking error (distance of virtual to true tool tip in video) for S2d, was observed to be 5 mm (mean), with a maximum of 10 mm.

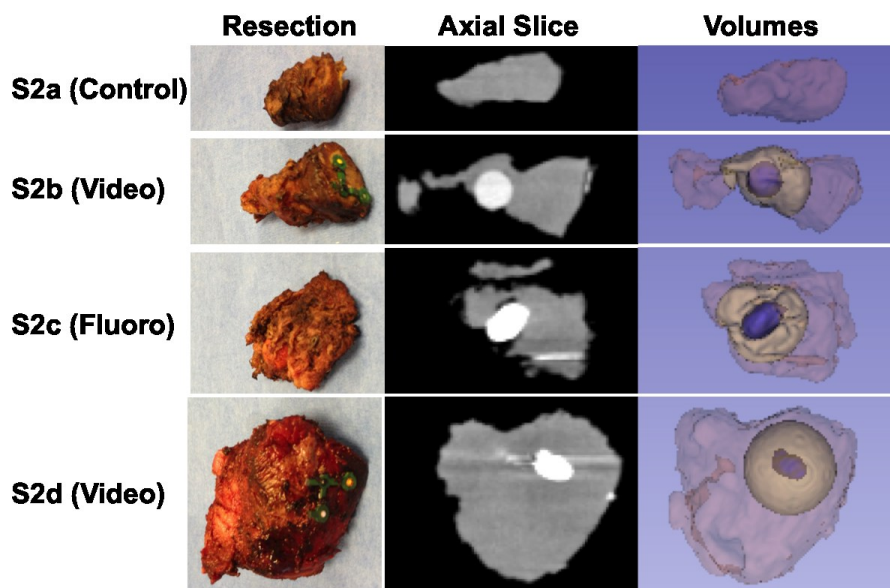


Figure 6.6 Resected tumor and margins with corresponding slices (tumor in high intensity, white) and segmented volumes (tumor in blue, intersection with ideal margin in yellow) from postoperative CBCT.

TABLE 6.2 TORS EXPERIMENTAL RESULTS

		Specimen	Margin	Tumor	Margin/	Margin/ Ideal	Tumor/	Tumor/
Experiment		[mm^3]	[mm^3]	[mm^3]	Specimen	Margin	Margin	Ideal Tumor
Set 1	S1a	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	S1b	8059.67	2960.45	257.99	0.37	0.71	0.09	0.49
	S1c	19602.55	1832.28	280.17	0.09	0.44	0.15	0.54
	S1d	13450.55	3410.59	583.45	0.25	0.81	0.17	1.11
Set 2	S2a	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	S2b	5485.52	3646.92	307.47	0.66	0.87	0.08	0.59
	S2c	6656.81	2429.50	323.08	0.36	0.58	0.13	0.62
	S2d	22287.22	4199.81	220.32	0.19	1.00	0.05	0.42
Ideal Margin: A sphere of radius = 10 mm					Ideal Tumor: A sphere of radius = 5 mm			

#### 6.1.2.5 Discussion

This is a proof-of-concept study that assessed the value of augmenting the surgeon's endoscopic view with CBCTA data with the goal of improving surgical accuracy and optimizing margins. Though limited by a single experienced TORS surgeon performing the resections, results demonstrate the value of video augmentation by improved margin status.

In these experiments, video-based augmentation (S1b, S2b, S1d, S2d) achieves superior tumor resection compared to fluoroscopy-based guidance (S1c, S2c). The improvements by S2d compared to S2c can be attributed to the supplementary guidance with margin delineation. However, superior results achieved by S1 scenarios, where overlays did not include margins, emphasize a disadvantage for 2D fluoroscopy, compared to 3D video augmentation, and the significance of the method of integration between supplemental navigational information and the primary visual field. Unlike the proposed video augmentation system, the fluoroscopic overlays are rendered through TilePro<sup>®</sup> (i.e., visible in the SCC), and shown below the native stereoscopic view port. Informal surveys and similar work for monocular video augmentation in skull base surgery<sup>98</sup> have suggested advantages of guidance through augmentation<sup>268</sup> of the primary

“natural” window. Improvements from S1c to S2c support the need to be able to enlarge regions of interest in order to take further advantage of the sub-millimeter resolution of 2D X-rays. The EV phantoms present an abnormally challenging environment consisting of a featureless tongue volume. As the IV resection proceeded, dental, oropharyngeal, and neurovascular anatomies serve as landmarks, while our simulated control EV models only provide superficial features (surface fiducials).

Despite encouraging results achieved by the proposed video augmentation system, issues of robustness and accuracy remain. Video augmentation registered initially is reliable on approach; however, during intraoperative resection, overlaid models should be updated to reflect surgical deformations. Intraoperative resected volume updates based on custom fiducials was susceptible to failure when the fiducial was not positioned orthogonally to the endoscope. For improvements on robustness, future work should look to incorporate Kalman filters and prior state information from tool tracking. In addition, a 5 mm (mean) TTE, not acceptable for TORS applications, can be improved through intraoperative fluoroscopy, using 2D3D registration to correct for kinematic inaccuracies, tissue deformation, and external forces.

## **6.2 Cardiothoracic Robotic Surgery**

This section summarizes a modified reproduction of “A Pilot Study of Augmented Reality from Intraoperative CBCT for Image-Guided Thoracic Robotic Surgery” as presented by Liu *et al.*<sup>269\*\*</sup> at the 2014 Hamlyn Symposium on Medical Robotics. This

---

<sup>\*\*</sup> The development of the pilot study and a system integrating intraoperative CBCTA for guided thoracic robotic surgery was completed by the author in collaboration with Drs. Mahdi Azizian, Jonathan M. Sorger, Benedetto Mungo, Oliver Wagner, Daniela Molena, and Russell H. Taylor. Drs. Wagner and Mungo

effort is a feasibility study using augmented reality from preoperative imaging and our proposed system for cardiothoracic robotic surgery with an *in vivo* ovine model.

### **6.2.1 A Pilot Study of Augmented Reality from Intraoperative CBCT for Image-Guided Thoracic Robotic Surgery**

Surgical approaches to resect lung tumors within a wedge, segment, or lobe are derived from preoperative volumetric data (e.g., from computed tomography (CT)). In a standard clinical setting, the image data is obtained with the lung inflated, ideally in end-inspiratory hold, and the patient in a supine position. Intraoperatively, the patient is rotated laterally, and the lung is collapsed in order to create a suitable surgical workspace in the thoracic cavity. Correlating information from preoperative datasets to the surgical scene is currently a challenging cognitive correlation; thus, the accuracy of these practices not only varies according to the surgeon's experience but is also subject to inconsistencies.

#### **6.2.1.1 Introduction**

This section presents an initial pilot study to evaluate the applicability of augmented reality using intraoperative cone-beam CT (CBCT) with the *da Vinci*<sup>®</sup> *Si* system for a pulmonary wedge resection. Augmentation of thoracoscopic video with planning data defining the target and critical structures offers the potential to provide initial localization of the pulmonary segment of interest. Although collapsing the lung

---

provided clinical consultation. Dr. Molena served as the clinical principal investigator and served as the surgeon participating in the experiments. Drs. Azizian and Sorger provided technical and academic expertise in system engineering and experimental protocol. Dr. Taylor served as the academic principal investigator for the work presented.

deforms the bronchi and parenchyma, we explore whether relative deformation of select mediastinal structures may still provide adequate intraoperative overlay. Using a porcine model, we measured deformation from intraoperative deflation and patient position. Furthermore, using an *in-vivo* ovine model, we tested the proposed image-guidance system by conducting a robotic wedge resection, guided by video augmentation from intraoperative CBCT.

### **6.2.1.2 Material and Methods**

#### **I. Evaluation of Lung Deformation**

In standard clinical practice, a preoperative CT image is acquired with the patient supine and lungs inflated. Intraoperatively, the patient is rotated laterally (surgical side up), and the lung is collapsed. An imaging study was performed on a live pig in order to estimate lung deformation from both intraoperative deflation and patient rotation. Three mock tumors were created with urethane, medium durometer spherical medical balloons (10 mm diameter), then filled with 0.5 ml of rigid polyurethane foam and 0.3 ml of acrylic paint. The tumor targets, along with a peri-tumor metal fiducial (52100 Chromium 1 mm diameter spheres), were placed in the animal's left lung using a radiopaque FEP I.V. catheter (Abbocath®-T 14G x140 mm). We imaged the porcine thorax using volumetric CBCT (Siemens syngo DynaCT, 90 kVp, 290 mA, 0.48x0.48x0.48 mm<sup>3</sup> voxel size) in the following four positions: 1) *Supine, Inflated (SI)*, 2) *Supine Collapsed (SC)*, 3) *Lateral Inflated (LI)*, 4) *Lateral Collapsed (LC)* (Figure 6.7). We segment the three peri-tumor fiducials, pulmonary bifurcation, and hemiazygos vein (proximal to the medial tumor) in each of the four data sets as targets. We compute the vector of each target created from inflated to collapsed. In the same patient position (Supine/Lateral), we

measure the absolute value of the vector; however, for data differing in rotation, we compared their L2-norm.

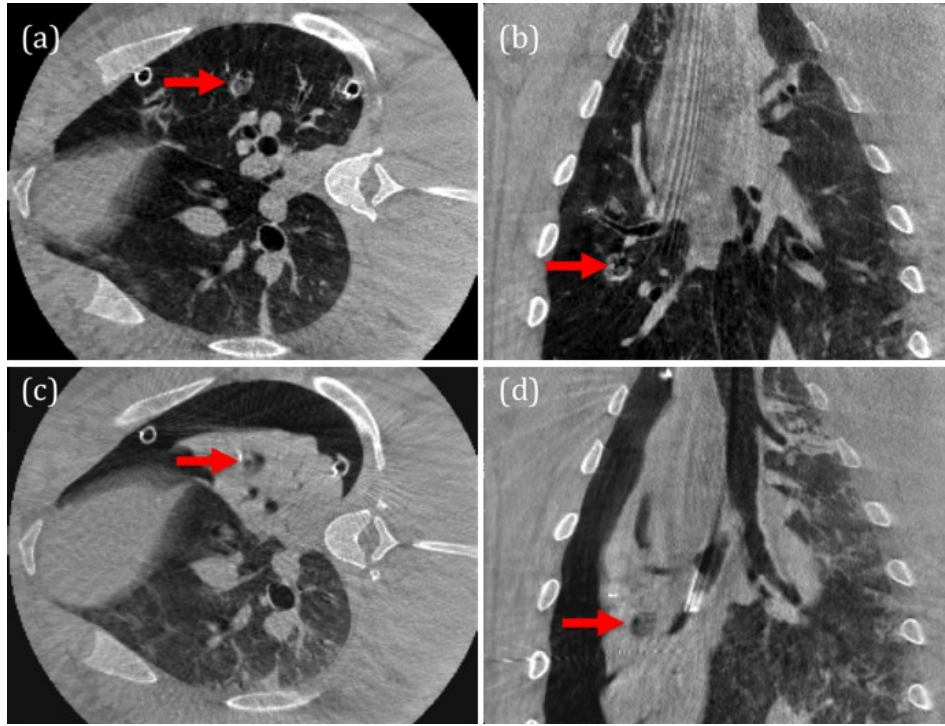


Figure 6.7 (a) Axial and (b) coronal slice of lateral inflated lung. (c) Axial and (d) coronal slice of lateral collapsed lung with a medial mock tumor indicated by red arrow.

## II. *In Vivo* Robotic Experiment

A robotic wedge resection in an ovine model was conducted to evaluate the feasibility and clinical workflow required for augmented reality from intraoperative CBCT. A mock tumor and a peri-tumor fiducial were placed in the left upper lobe using the same protocol as above. The ovine model was placed in a right lateral position. We then collapsed the lung and used manual laparoscopy to suture five plastic registration fiducials onto the abdominal wall before re-inflation. A CBCT acquired at this step not only captured intraoperative patient position, but also images the current physiology and

anatomy, which may differ from preoperative scans obtained in clinic. Next, segmentation of the critical anatomy of interest (i.e., the tumor, pulmonary artery, hemiazygos vein) and registration fiducials is completed manually (Figure 6.8c). Automatic segmentation in medical imaging is an extensive field of research which may be potentially addressed in future work. The segmented critical anatomy is registered to the stereoscopic video using the 3D interface of the video augmentation software by manually aligning the registration fiducials and adjusting with visible anatomical landmarks (Figure 6.8a). This augmentation system, described in detail elsewhere<sup>266</sup>, is implemented by extending *cisst/SAW*<sup>129</sup>, open source libraries supported by the ERC/LCSR center at the Johns Hopkins University. The augmented scene overlays the tumor, pulmonary artery, and hemiazygos vein onto the stereo endoscopic images (Figure 6.8a, b, c). A fellowship-trained cardiothoracic surgeon<sup>††</sup> with specific expertise in minimally invasive interventions is tasked to remove the synthetic tumor with a wedge resection (Figure 6.8d).

---

<sup>††</sup> Dr. Daniela Molena, Director, Robotic and Minimally Invasive Thoracic Surgery and Assistant Professor of Surgery, Johns Hopkins Hospital, Baltimore, MD.

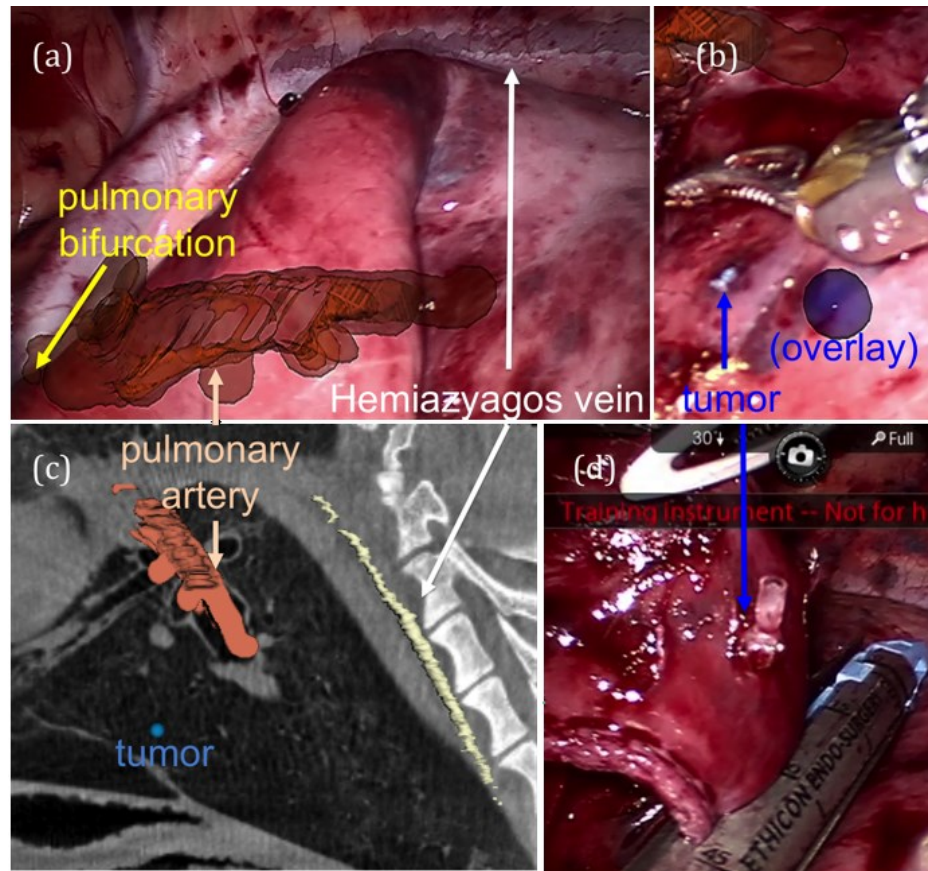


Figure 6.8 Overlays from anatomic registration of (a) pulmonary landmarks and (b) mock tumor (c) Segmentation of critical structures in CBCT (sagittal slice of lateral, inflated data). (d) Wedge resection.

### 6.2.1.3 Results

Differences of vector norm between targets from standard preoperative image acquisition (SI) to intraoperative setup (Figure 6.7 c, d) are measured at  $19.6 \pm 11.4$  mm, while changes from a lateral rotation average  $9.0 \pm 8.9$  mm. Deformation solely from lung deflation was measured at  $20.1 \pm 9.2$  mm and  $13.9 \pm 7.3$  mm for supine and lateral positions, respectively. Other studies<sup>270</sup> tracking surface points (compared to our targets embedded in parenchyma) have reported lung (porcine) deflation displacement ~48-63 mm. In fact, the displacement from deflation of natural landmarks used in registration



(i.e., pulmonary bifurcation and hemiazygos vein) measured  $3.3 \pm 1.4$  mm. From postoperative analysis of recorded video, augmentation of the tumor (LI) was optically measured at  $\sim 25$  mm away from the actual target. The robotic experiment successfully removed the entire tumor within a wedge resection, with results confirmed visually through post-experimental dissection of the resected specimen, as well as fluoroscopically using an X-ray indicating the presence of the metal peri-tumor fiducial within the wedge.

### **6.2.2 Discussion**

Experimental results from an *in-vivo* ovine pilot study demonstrated the capability of augmented reality with intraoperative CBCT to provide initial localization for thoracic robotic wedge resection. A 3.3 mm deformation for mediastinal structures is adequate for registration purposes. Overlays of embedded parenchyma targets from inflated acquisitions onto collapsed lung may be displaced  $\geq 20$  mm, but can still provide navigational value, especially for inexperienced surgeons. For other thoracic interventions involving vascular dissection (e.g., lobectomy and segmentectomies), systematic accuracy must be improved. This was an initial pilot study and further studies are needed to evaluate quantitative outcomes. Future work should explore overlays from multi-modal preoperative volumetric data, automated segmentation, the precedence of natural landmarks, and deformable intraoperative registration.

## 6.3 Endoscopic Skull Base Surgery

This section summarizes a modified reproduction of Liu *et al.*<sup>98††</sup>, originally presented at the 2012 Proceeding of SPIE Medical Imaging. We obtained approval from the Institutional Review Board (IRB) at the Johns Hopkins Hospital to conduct a pilot clinical study on video augmentation for image-guided endoscopic endonasal skull base surgery<sup>§§</sup>. The purpose of the study is to assess feasibility and benefit to surgical performance from overlaying CT or MR planning data in real-time, high-definition endoscopic video. We describe our efforts in translating our system from development in a research laboratory into a form suitable for a clinical arena.

### 6.3.1 A Clinical Pilot Study of a Modular Video-CT Augmentation System for Image-Guided Skull Base Surgery

The described clinical pilot study aimed to evaluate fifteen patients undergoing skull base tumor surgery in which each surgery includes the experimental video-CT system deployed in parallel to the standard-of-care (un-augmented) video display. Preoperative planning included segmentation of the carotid arteries, optic nerves, and surgical target volume (e.g. tumor). An automated camera calibration process is developed that demonstrates mean re-projection accuracy of  $(0.7 \pm 0.3)$  pixels and mean

---

†† The development of the pilot study and clinical translation of the Video-CT system was completed by the author in collaboration with Drs. Daniel J. Mirota, Mr. Ali Uneri, Drs. Yoshito Otake, Gregory Hager, Douglas D. Reh, Masaru Ishii, Gary L. Gallia, and Jeffrey H. Siewerdsen. Dr. Gallia was the clinical principal investigator while Dr. Siewerdsen served as the academic principal investigator. The cisst/SAW infrastructure was developed by Dr. Peter Kazanzides, Anton Deguet, Balazs Vagvolgyi and others. The TREK infrastructure was developed by Ali Uneri and others.

§§ After a single patient evaluation, as of June 2014 this IRB study, has been suspended due to technical barriers unrelated to our image guidance system.

target registration error of  $(2.3 \pm 1.5)$  mm. Questionnaires distributed to clinicians (one neurosurgeon and two otolaryngologists) are used to assess primary outcome measures regarding the benefit to surgical confidence in localizing critical structures and targets by means of video overlay during surgical approach, resection, and reconstruction.

#### **6.3.1.1 Introduction**

Endoscopic endonasal skull base surgery is an emerging minimally invasive approach used to address a broad spectrum of skull base lesions. Precise visualization is necessary to ensure complete resection within the complex anatomy of the endonasal space<sup>32</sup>. Such skull base pathologies are in close proximity to critical neurovascular structures, and encroachment can have significant consequences (e.g., neurological injury and death). Ongoing research to improve skull base surgery guidance includes virtual endoscopy<sup>271</sup>, image-overlay<sup>272,273</sup>, and intraoperative cone-beam CT<sup>230,274</sup>. Improved visualization using 3D endoscopes<sup>275</sup> also offers a novel technique to improve patient safety and reduce clinical learning curves. The system described below extends such work in a novel modular architecture for video augmentation that automates the camera calibration process and can be adapted to other endoscopic or laparoscopic procedures. The system includes a streamlined calibration process consistent with clinical workflow and provides registration of the video scene with preoperative (or intraoperative) image data (e.g. structures defined on images from CT or MRI).

#### **6.3.1.2 System Architecture**

The video-CT system extends the TREK<sup>126</sup> software architecture (Figure 6.9) for image-guided surgery to a clinically practical form. As described previously, TREK binds open-source libraries for image visualization and analysis from 3D Slicer<sup>132</sup> (na-mic

kit, Brigham & Women's Hospital, Cambridge MA) and real-time tracking and registration from the *cisst* libraries<sup>129</sup> (ERC, Johns Hopkins University, Baltimore MD). These modular components of the framework are illustrated in Section 3.3. The front-end graphical user interface loads preoperative CT or MRI images and corresponding planning/segmentation data after processing with ITK-Snap<sup>244</sup>. We extended the *cisst* package (specifically the computer vision, device interface, and tracking functionalities) to include an automatic camera calibration and hand-eye calibration described below. Interface to a clinical tracking system (StealthStation, Medtronic Inc., Minneapolis MN) provide infrared tracking of a rigid body marker attached to the endoscope and a reference marker attached to the stereotactic head frame. All studies involved a high-definition (HD) video endoscope (H3-Z Camera, Karl Storz Inc., Tuttlingen Germany).

### **6.3.1.3 Surgical Planning**

Excision of skull base tumors is a challenge even for experienced surgeons for numerous reasons, including the proximity of surgical targets to critical anatomy, such as the carotid arteries and cranial nerves. Preoperative diagnostic imaging, including CT, CT angiography (CTA), and MRI, provides a wealth of 3D anatomical information of these areas of interest. However, a conventional intraoperative guidance system involves the separate unregistered display of such 3D images apart from the endoscopic video. Within such preoperative image data the surgical approach (trajectories), as well as segmentation of pertinent anatomical structures, the surgical target, and margins - referred to simply as "planning data" can be defined. The video augmentation workflow (Figure 6.10) described below include an offline, preoperative process to define such planning data, specifically to segment critical structures (i.e., the carotid arteries and optic nerves) and

the surgical target (i.e., the tumor volume) (Figure 6.11a). Using ITK-Snap (NLM Insight Toolkit, University of Pennsylvania, Philadelphia PA), these structures are defined in preoperative CT/CTA and/or MR using semi-automatic region growing and thresholding complemented by manual refinement and final review by the operating surgeon.

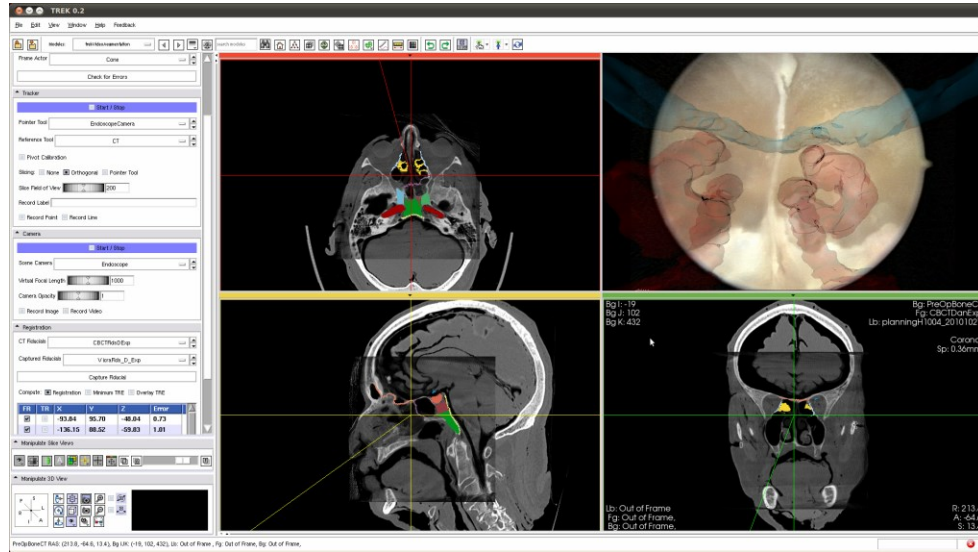


Figure 6.9 Screenshot of the video-CT system's interface including the augmented display and triplanar view of the CT and surgical plan.

### 6.3.1.4 Video-CT Registration

#### 6.3.1.4.1 Camera Calibration

Translating the system from a research platform to a clinically useful form necessitated a fast process for calibration and registration of the endoscopic camera. Please refer to Section 3.5.1 for details regarding hand-eye calibration used in this work.

#### 6.3.1.4.2 Hand-eye Calibration

Please refer to Section 3.5.2 for details regarding camera calibration used in this work.

#### **6.3.1.5 Preclinical Tests: Cadaver Studies**

Having completed the technical development required for the translation of the research system to be clinically deployable, the system workflow and usability are tested in a series of preclinical cadaver studies. These sessions focused on not only feasibility and quality assessment of the system but also provided opportunities to familiarize the surgeon with the video-CT user interface. Care was taken to follow clinical steps and setup as closely as possible to accurately assess the workflow (Figure 6.10).

Cadaver head specimens were imaged with standard CT, and critical structures were segmented with ITK-Snap. Similar to the target and critical anatomy anticipated in skull base surgery, mesh segmentations were created for the carotid arteries, optic nerves, and pituitary gland as shown in Figure 6.12 in red, blue, and magenta, respectively. Such segmentation was performed offline as a preoperative step by a trained technologist, and the resulting "Surgical Plan" was reviewed and refined if necessary by the operating surgeon. The datasets were loaded onto a StealthStation and further processed for the RAS to image space vector (IJK) that aligns the CT coordinate system in TREK video-CT to coordinates from the StealthStation patient registration. Following calibration and registration of the endoscope and other tracked tools, patient registration was conducted according to the standard-of-care with the StealthStation, completing the setup process for the operation.

Figure 6.12 illustrates three of the views available in the guidance system, namely triplanar views (CT or cone-beam CT) overlaid with planning data along with real-time endoscopic video (without or with planning data overlay). Figure 6.11a shows a sagittal slice of the CT image with planning data superimposed, which is optionally displayed in

the video-CT user interface along with other triplanar views. Figure 6.11b shows an endoscopic view in the region of the sphenoid sinus as in the standard-of-care (un-augmented) video, whereas Figure 6.11c shows the same HD video scene augmented in real-time with overlay of the carotid arteries. Certain parameters of the video-CT system are exposed to the user for customizing the visualization (e.g., adjustment of the color palette, lighting, opacity of critical structures, and placement of the virtual camera and focal lengths).

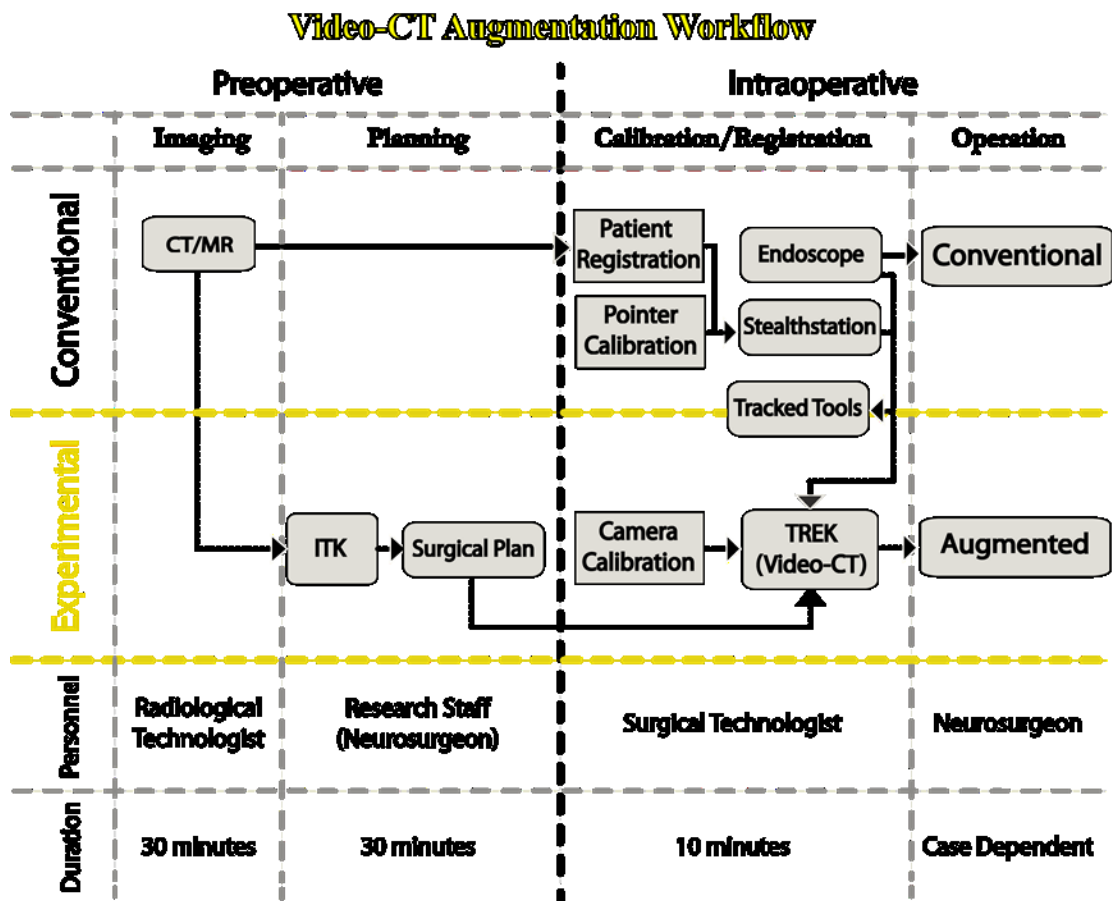


Figure 6.10 Video-CT Augmentation workflow for image-guided endoscopic endonasal surgery.

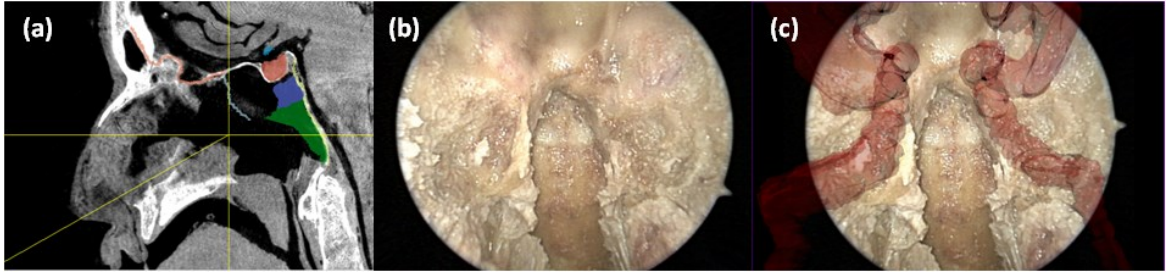


Figure 6.11 (a) Example planning data shown in a sagittal CT slice of a cadaver employed in preclinical evaluation. Target structures include the anterior skull base (pink), the inferior clivus (green), the superior clivus (blue), and the surgical target (pituitary, in red). The clinical pilot study involves side-by-side display of (b) conventional (non-augmented) endoscopic video and (c) the experimental system for augmentation and real-time overlay of registered planning data within the video scene. The carotid arteries, which in this case have a narrow inter-carotid distance, are overlaid in red in a trans-sphenoid clival drillout procedure.



Figure 6.12 (a)-(c) Three perspectives of video-CT overlay as the endoscope is panned left-to-right and anterior-to-posterior in a preclinical evaluation following a cadaveric clival drillout. The images show overlay of the carotid arteries (red), optic nerves (blue), and pituitary gland (purple) in trans-sphenoid approach to the skull base.

#### 6.3.1.6 Clinical IRB Case Study



The video-CT augmentation system was deployed in an IRB-approved clinical pilot study<sup>\*\*\*</sup> in parallel to a conventional standard-of-care (un-augmented) endoscopic video display, as shown in Figure 6.13. The pilot study looked to recruit ~15 neurosurgical patients and assess primary outcome measures focusing on expert assessment of the utility of the video-CT overlay, the potential benefit to surgical confidence, and the visualization of critical structures on specifically delineated phases of approach/exposure, resection, and reconstruction by three surgeons (1 neurosurgeon and 2 otolaryngologists). Following each case, both the neurosurgeon and otolaryngologist<sup>†††</sup> are asked to respond to the questionnaire summarized in Table 6.3. Questions #1-3 provide ordinal ratings (score 1-5) by the following utility scale: 1 = Significant hindrance / Negative effect; 2 = Minor hindrance / Slightly negative effect; 3 = Not helpful / No benefit or hindrance; 4 = Somewhat helpful / Slight benefit; 5 = Very helpful / Major benefit. Questions #4-5 allow free response in relation to anatomical and disease variations outside those present within the particular case that would potentially benefit from video-CT overlay.

The potential benefit to surgical confidence and visualization of critical structures is clear: although the carotids are evident in varying degrees to a trained surgeon in protuberances based on bony landmarks and color variations on the posterior aspect of

---

<sup>\*\*\*</sup> After a single patient evaluation, as of June 2014 this IRB study, has been suspended due to technical barriers unrelated to our image guidance system.

<sup>†††</sup> Dr. Gary L. Gallia, Assistant Professor of Neurosurgery and Oncology, Director of Endoscopic and Minimally Invasive Neurosurgery, Director of the Neurosurgery Skull Base Center, Johns Hopkins Hospital, Baltimore, MD. Dr. Douglas D. Reh, Medical Director, Otolaryngology- Head and Neck Surgery at Johns Hopkins Hospital, Lutherville, MD. Dr. Masaru Ishii, Associate Professor of Otolaryngology- Head and Neck Surgery, Johns Hopkins Hospital, Baltimore, MD.

the sphenoid sinus, the augmented video-CT display can provide a visually significant improvement in the conspicuity of such subtleties, particularly in the context of a bloody surgical field, in cases exhibiting anatomical variations, and in situations where the artery is encased by tumor<sup>276</sup>. The primary outcome of the clinical study is to record subjective assessment via the questionnaire summarized in Table 6.3, with analysis of the operator responses pending completion of the pilot study. Assessments included evaluating the video-CT augmentation in context to improve surgical confidence, localization accuracy, and efficiency (Question #1 in the questionnaire). The utility of the video-CT system in improving visualization of these critical structures during the approach, surgical resection, and reconstruction is assessed in Questions #2 and #3 of the questionnaire.

TABLE 6.3 SUMMARY QUESTIONNAIRE FOR EXPERT ASSESSMENT OF VIDEO-CT IN SKULL

BASE SURGERY

<b>Postoperative Questionnaire</b>	
<b>1. Rate the effect of video-CT overlay on overall:</b>	
	Surgical confidence {1-5}
	Certainty in structure localization {1-5}
	Efficiency {1-5}
<b>2. Rate the utility of video-CT overlay in visualizing:</b>	
	Carotid arteries {1-5}
	Optic nerves {1-5}
	Surgical target (specify) {1-5}
<b>3. Rate the utility of video-CT overlay during:</b>	
	Surgical approach {1-5}
	Target resection {1-5}
	Reconstruction {1-5}
<b>4. What variations in <u>normal anatomy</u> pertinent to this case would benefit the most from video-CT overlay?</b>	
<b>5. What variations in <u>pathology</u> pertinent to this case would benefit the most from video-CT overlay?</b>	

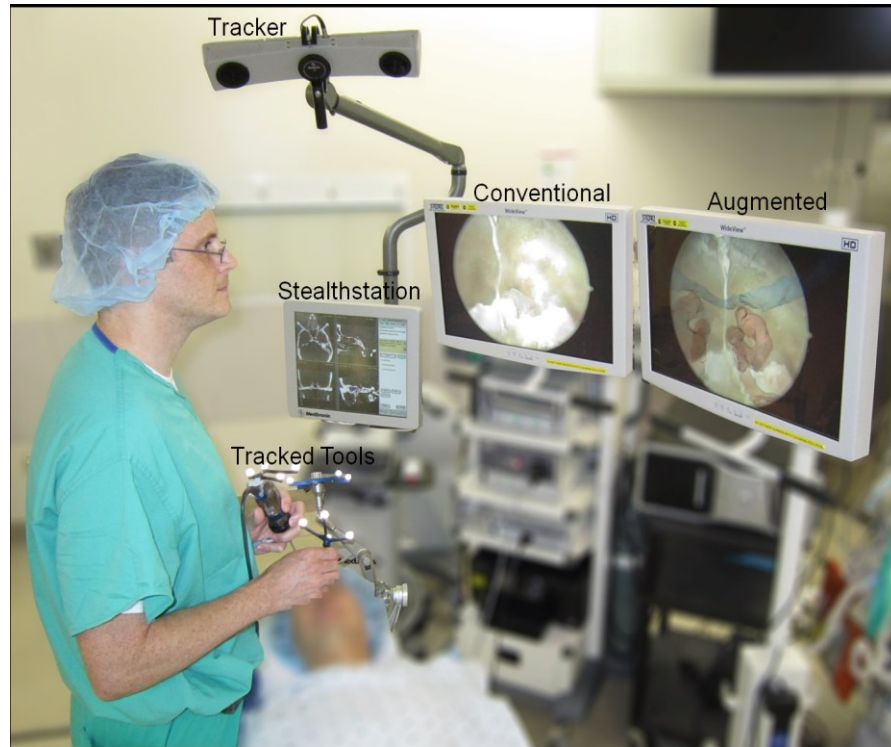


Figure 6.13 Clinical study operating room setup, showing components from the standard-of-care and video augmentation system.

### 6.3.2 Discussion

This work describes the development and translation of a video-CT augmentation system for an endoscopic skull base surgical clinical study. The registration accuracy of the current system is limited by the accuracy of the optical tracking system, but has potential future improvements through the incorporation of intraoperative C-arm cone-beam CT and 3D image-based registration<sup>97</sup>. Automation of the camera calibration process streamlines the video-CT registration system to a form suitable for use by a trained OR technologist in a manner that is consistent with surgical workflow. The clinical pilot study described herein provides a valuable basis for use of video-CT

registration in routine clinical care and critically evaluates the utility and cases for which such capability will have most benefit.

## **6.4 Robotically-Assisted Cochlear Implant**

This section presents a modified reproduction of Liu *et al.*<sup>277\*\*\*</sup> published in JAMA Otolaryngology Head & Neck in 2014. To the best of our knowledge, this is the first reported study for robot-assisted cochleostomy and mastoidectomy using a *da Vinci® Si system* with augmented reality. Two cadaveric case studies were conducted using the *da Vinci®* system to test the feasibility of a clinical workflow for robot-assisted cochleostomy. The second case extends the first with integration of CBCT-based image guidance through stereo video augmentation of segmented critical anatomical structures.

### **6.4.1 Cadaveric Feasibility Study of da Vinci Si–Assisted Cochlear Implant With Augmented Visual Navigation for Otologic Surgery**

Confidence in precision and localization is critical for otolaryngology - head and neck surgeries, including the surgical placement of a cochlear implant. Registration of the preoperative/intraoperative image data can map significant neurovascular structures and target trajectories (derived from standard preoperative computed tomography (CT)/cone-beam CT (CBCT)/MRI images) directly to the patient's operative workspace. In fact,

---

\*\*\* The development of the feasibility study and augmented visual navigation for guided otologic robotic surgery was completed by the author in collaboration with Drs. Mahdi Azizian, Jonathan M. Sorger, Russell H. Taylor, Brian K. Reilly, Kevin Cleary, and Diego Preciado. Drs. Azizian and Sorger provided technical and academic expertise in system engineering and experimental protocol. Dr. Taylor was an academic collaborator in this effort where Dr. Cleary, Children's National Medical Center, Washington, D.C., served as the academic principal investigator, particularly in the design of the tool adapter used to attach the clinical drill. Dr. Preciado, an otolaryngologist at the Children's National Medical Center, Washington D.C served as the clinical principal investigator while consulting with Dr. Kelly.

head and neck surgical literature reports the usage of optically tracked pointers (Medtronic StealthStation®) for image-guided endoscopic endonasal skull base surgery<sup>97,98,257</sup>. A *da Vinci*® *Si* system is used in bilateral cadaveric cochleostomies and mastoidectomies (Case 1: patient left, robot-assisted; Case 2: patient right, robot-assisted with augmented reality). A custom 3D printed drill adapter (Figure 6.14 inset) allowed a standard manual drill to be attached with a 30 degree offset on the shaft of an 8 mm *da Vinci*® *Si* tool.

#### **6.4.1.1 Introduction**

For cochlear implants (CI), the accuracy of the mastoidectomy, cochleostomy, and insertion angle of the device has been shown<sup>35-37</sup> to be critical for device function and clinical functional outcomes. Novel technology integrating preoperative/intraoperative image data has the potential to significantly improve the accuracy of the implantation approach through improved facial recess widening and cochleostomy placement. Image guidance can be done through registration of standard preoperative computed tomography and/or cone-beam computed tomography (CBCT) images coupled with mapping of the fixed temporal bone landmarks and target trajectories directly to the patient's operative workspace. Other groups<sup>278</sup> have described a stereotactic frame carefully mounted to the patient's skull and surrounding mastoid, which facilitates a surgical drilling trajectory that enables a percutaneous approach to implantation. In fact, the use of optical tracking systems (e.g., StealthStation; Medtronic) for image-guided endoscopic endonasal skull base surgery is well described<sup>97,98,257</sup> in clinical use and robot-assisted otoneurosurgery<sup>279-283</sup>, as well as computer-assisted implant placement for CI surgery<sup>40,280,284-287</sup>, and is an ongoing active area of research.

The goal of this study is to test whether a well-developed and industry-standard system (*da Vinci* Surgical System; Intuitive Surgical, Inc.) could facilitate cochlear implant. The study uses a *da Vinci Si* system with custom tool adapters for master-slave-assisted cortical mastoidectomy, posterior tympanostomy, and cochleostomy. Two cadaveric case studies were conducted using the master-slave system to test the feasibility of a proposed clinical workflow. In the second case, we introduced additional steps incorporating intraoperative CBCT-based image guidance through stereo video augmentation with direct overlay of critical anatomy (facial nerve, cochlea, and round window).

#### **6.4.1.2 Methods**

A *da Vinci® Si* system is used in bilateral cadaveric cochleostomies and mastoidectomies. Institutional review board review from Children's National Medical Center was waived for this study. Case 1 was a donor cadaver left temporal bone, with the CI surgery being performed with the master-slave system. Case 2 was a donor cadaver right temporal bone, with the master-slave system assisted with augmented reality. To dissect the temporal bone, a custom drill adapter was fabricated using a 3-dimensional (3D) printer (Objet Eden500V; Stratasys, Ltd). This secured the attachment of an osteon pneumatic drill (CONMED) with a 30° offset on the shaft of an 8-mm *da Vinci Si* tool. The offset strategically positioned the *da Vinci Si* arm away from the workspace while allowing the drill shaft effective parallelization with the axis of the endoscope. High definition (1080i @ 60Hz) 3D visualization, rendered through the surgeon's console, using a 12 mm (0°) *da Vinci® Si* endoscope had a minimal pixel resolution of 0.4 mm/pixel at a corresponding near field measurement of 7 mm and

provided appropriate viewing for the surgeon. Digital magnification ranged from 1x-4x (cochleostomy), while the scale of the master-to-slave manipulators was fixed at a ratio of 3-to-1.

The workflow steps involving cases 1 and 2 are outlined in below. The second case extended the workflow steps of the first with integration of CBCT-based image guidance through stereo video augmentation of segmented critical anatomic structures

Soft-tissue work to expose the mastoid is performed without the *da Vinci* Si. The surgeon exposed the mastoid bone and retracted the ear forward. The following workflow was used for Case 2, and the workflow for Case 1 included all of the components listed below, except for those identified with an asterisk.

- 1.\*Expose mastoid bone and retract ear forward to fix three self-drilling zinc pan head Phillips screws 6mm long (No. 6 drill bit used to initiate entry; custom drill adapter shown in Figure 6.14 inset), placed with their centroid above the inner ear to an exposed right mastoid (Figure 6.15a).

- 2.\* Acquire preoperative CBCT scan (Siemens Powermobil; Siemens; head protocol, 109 kilovolt [peak], 290 mA, 0.48x0.48x0.48 mm<sup>3</sup> voxel size) with the isocenter positioned around the cochlea.

- 3.\* Use ITK-Snap (<http://www.itksnap.org/>)<sup>244</sup> to manually segment the facial nerve, basal turn of the cochlea, and the round window (Figure 6.15b), as well as the divots of 3 fiducial screw heads from preoperative CBCT.

4. Position the components of the *da Vinci Si* as shown in Figure 6.14, with the base of the patient side cart on the opposing side of the patient target.
5. Use the printed adapter to attach the drill at a 30° offset onto the shaft of an 8-mm *da Vinci Si* tool that is inserted into the primary robotic arm while either a *da Vinci Si* suction or irrigator tool (EndoWrist; Intuitive Surgical, Inc) is placed or a standard suction/irrigator device is fixed to the secondary arm.
- 6.\* Conduct manual point-based registration of endoscopic view (video) to the CBCT<sup>266</sup> by identifying the fiducial screws.
7. Perform mastoidectomies, posterior tympanotomies, and cochleostomies using the *da Vinci Si*.
- 8.\* Perform manual refinement of cochleostomy and insertion of an implant wire phantom into the cochlea at the basal turn (Figure 6.16).





Figure 6.14 Layout of the operating room with the *da Vinci Si* for Case 1. Inset is a close-up of the initial position of the endoscope, suction/irrigator, and drill attached with the custom tool adapter.

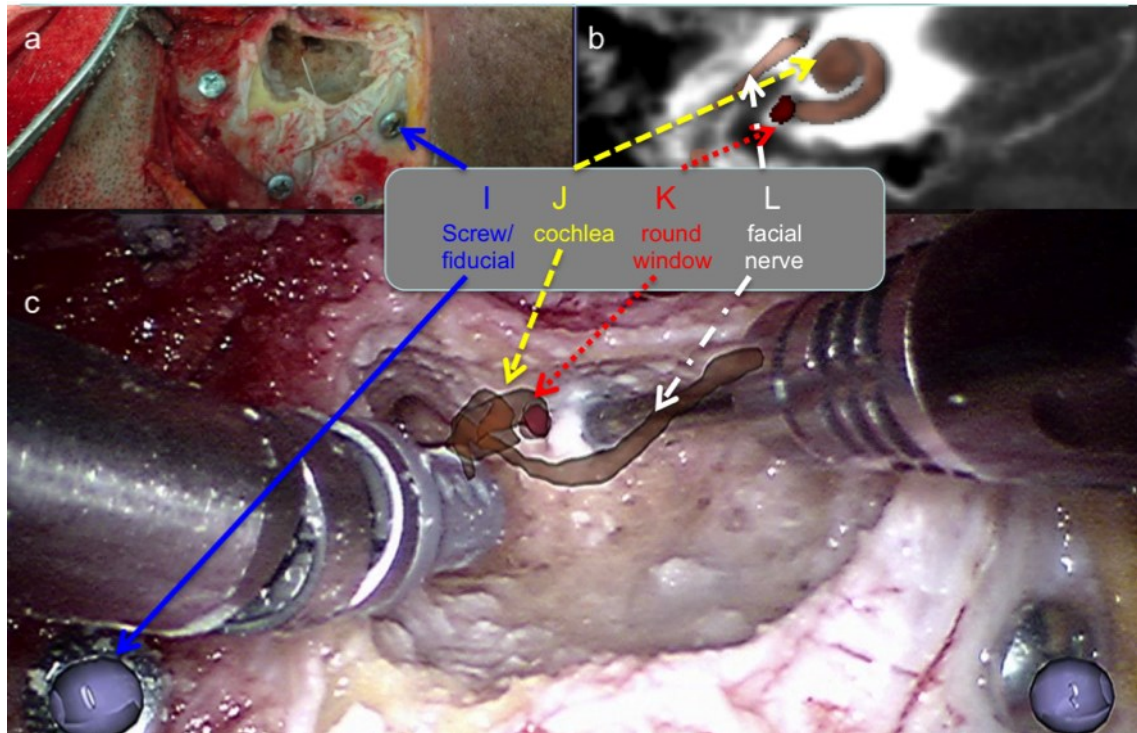


Figure 6.15 (a) Master-slave-assisted mastoidectomy and cochleostomy on cadaveric right temporal bone completed with augmented reality, registered using three fiducials. (b) Coronal slice in preoperative cone-beam computed tomography showing segmentation of the critical structures. (c) Monocular screen capture of the right eye during cochleostomy with video augmentation of the segmented models.

### 6.4.1.3 Results

These two case studies demonstrate the feasibility of master-slave-assisted cochleostomy and mastoidectomy with augmented reality integrated as image guidance on a *da Vinci Si*. Setup for the system after exposing the soft tissues and placing the fiducials, which included capturing the CBCT images, segmenting the critical anatomic structures, and configuring the augmented reality capability, require approximately 110 minutes. The mean surgical time for surgery per side was 160 minutes. This surgical time

included the time required to bring the *da Vinci Si* system to the field, mount the tools and drill to the robotic arms, and perform all aspects of the drilling and electrode array insertion. More importantly, there was no violation of critical structures and, in both cases, there was full insertion of the phantom electrode array. Accuracy of the augmented reality overlay was confirmed on the second side by uncovering the mastoid segment of the facial nerve. The position of the nerve visually corresponded with the overlaid image. We did not assess the position of the phantom array within the cochlear compartments with histopathologic testing. Similarly, the angle of insertion and the position of the cochleostomy were not confirmed with histological analysis.

#### Case 1: Master-Slave–Assisted Cochleostomy

The *da Vinci Si*–assisted drilling of the mastoid bone was conducted on the cadaver’s left ear without breaching of the facial nerve to expose the middle ear. Successful cochleostomy, which was confirmed with postoperative CBCT, was also performed at 4x digital magnification with the *da Vinci Si* system.

#### Case 2: Image-Guided Master-Slave–Assisted Cochleostomy With Stereo Video Augmentation

Similar to Case 1, Case 2 achieved successful cochleostomy following a mastoidectomy on the right cadaveric specimen. Additionally, postoperative image data confirmed the successful placement of an implant wire phantom in the cochlea (Figure 6.16). Registration of critical data from CBCT to robotic video using 3D augmented

reality of segmented models of vital anatomic structures was also visually confirmed after exposure of the facial nerve. Each full surgical procedure is completed within an average duration of 160 minutes. The custom drill adapter, which has a 30° offset and is coupled to the working da Vinci Si robotic arm, allow for enough degrees of freedom to navigate around corners with the drill and to complete all of the necessary motions required for drilling the temporal bone in a fluid natural fashion, not unlike freehand drilling. There was no injury or surgical breach of the tegmen tympani, sigmoid sinus, or facial nerve. After the facial recess/posterior tympanostomy was opened, the 3D endoscope camera provided an adequate view of the middle ear structures, including the stapes and the round window niche. Because of the restrictions related to magnification (dependent on the distance of the camera tip), this view of the middle ear is thought to be inferior to the standard stereoscopic vision conferred by a binocular microscope. This issue can be addressed in future subsequent studies by adjusting scope selection and the camera focal length. Finally, new suction and drilling equipment need to be engineered. In the present case studies, the articulating suction irrigator was not small enough to insert through the recess. As a result, the surgeon had to physically attach and couple a 3F Frazier tip to a robotic side port instrument to suction in the middle ear. Another limitation was the size of the robotic arm tools (grasping forceps) that are not currently designed for otologic surgery in the middle ear, thereby making it difficult to complete middle ear work through the facial recess using the master-slave system. Manual refinement was needed to finalize the insertion of the phantom electrode. Despite these minor limitations, we were able, in both surgical cases, to complete the entire procedure, including the cochleostomy, without difficulty using the *da Vinci Si* system.

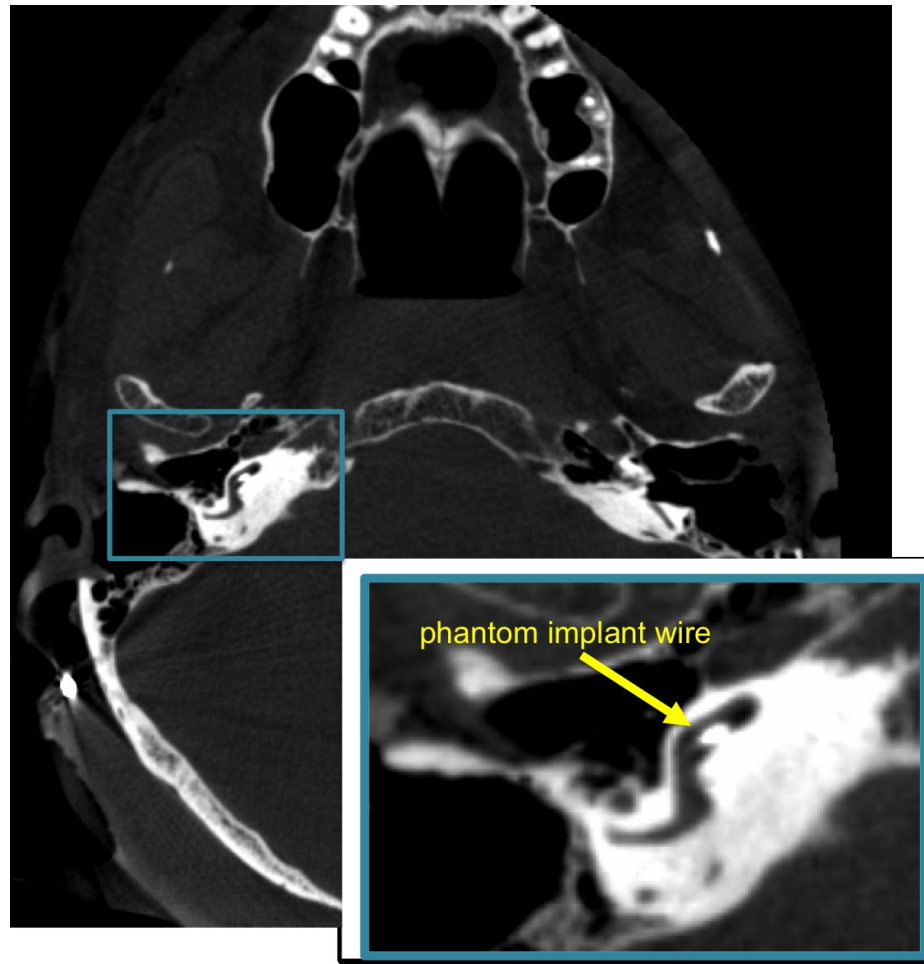


Figure 6.16 Axial slice from postoperative CBCT showing the successful placement of a phantom implant wire (yellow) in the cochlea.

#### **6.4.2 Discussion**

This initial cadaveric feasibility study shows that the master-slave approach with augmented reality is possible. The potential of this technology is vast. Indeed, if proved to be precise, reliable, and cost-effective, a surgical system that provides immediate intraoperative feedback of the anatomy would be quickly accepted by surgeons and patients, not only for cochlear implant but also for other neuro-otologic procedures, such as excision of cerebellopontine angle tumors, removal of petrous apex cholesterol

granuloma, labyrinthectomy, and decompression of the endolymphatic sac in Meniere's disease. Suboptimal placement of the electrode too inferior to the round window increases the likelihood of surgical damage to auditory spiral ganglion neurons with subsequent decreased hearing performance, as well as the potential for vestibular stimulation, cerebrospinal fluid leakage, and false placement in hypotympanic air cells, rendering the implant nonfunctional<sup>288</sup>. Although the incidence of each of these potential difficulties is rare, suboptimal placement of the electrode array is likely a contributing factor in cochlear implant soft failure rates, which can occur in up to 23% of implants<sup>289,290</sup>. Mitigating collateral forces resulting from electrode implant insertion on the cochlea is also a well-studied body of work<sup>291,292</sup>.

Research in computer-assisted approaches for cochlear implant includes custom robotic systems<sup>41,279</sup> and adaptations of industrial robots<sup>280,281,293</sup>. These groups have demonstrated the value of human-robot collaboration to filter hand tremor<sup>41</sup>, to support motion and force scaling<sup>282</sup>, and to avoid skipped movement and precise micro-drilling<sup>294,295</sup>. In the proposed approach using the *da Vinci Si* system, we take advantage of tool position stability as well as motion scaling and visual magnification. Other studies have focused on image guidance through path planning, which is especially significant for percutaneous approaches and implant electrode placement<sup>20,24,25</sup> with steerable arrays that provide force feedback. The approach described herein use augmented reality as a means of guidance. Rather than tracking a surgical tool to a registered radiologic image, the software allows the image of the critical structures to be "injected" and overlaid onto the live 3D endoscopic image in real time. To accomplish this, it is necessary to first segment the structures of interest from a CBCT scan, register to fiducials that are placed



in the surgical field of view, and then overlay a 3D image of the critical structures onto the actual 3D endoscopic view that the surgeon sees in the console. Similar to the use of optical navigation in sinus surgery, where image guidance has a more than 90% level of satisfaction with a reported accuracy of 1 to 3 mm in 79% of cases and only a few minutes needed for setup time<sup>296</sup>, the surgeon maintains direct visualization of the surgical workspace with enhanced navigation directly integrated to the primary field of view. Future extension to the image guidance used can include path planning<sup>40</sup> and virtual fixtures<sup>41-45</sup> to not only render critical structures but also improve safety and accuracy for surgical trajectories.

Undoubtedly, careful evaluation of the relative anatomy of the facial nerve, chorda tympani, orientation of the cochlear basal turn, round window anatomy, and cochlear axis based on in-room tomographic images is beneficial to cochlear implant surgeons to reduce risk, improve safety, and minimize time under anesthesia. In addition, precise control of the surgical drill through a robotic arm interface along with image augmentation in the microscope field-of-view increases the confidence of the surgeon, reduces operative risk and operative time, and minimizes the risk of postoperative complications. A major safety benefit of this study would be to mitigate facial paralysis with cochlear implant surgery, which has devastating psychological consequences and an estimated incidence as high as 1.1%<sup>297</sup>.

The *da Vinci* system has shown promise in surgery for many clinical indications, including prostate cancer, colorectal disease, renal disease, cardiac disease, and head and neck cancer. Although the *da Vinci* system has been primarily applied for handling soft tissue in anatomic regions that are difficult to access, the robotic arm has great potential

for lateral skull base procedures and, in particular, cochlear implants. Coupled with image guidance, we believe that the use of the *da Vinci* system has the potential to offer the surgeon enhanced surgical accuracy, allowing for more precise cochlear implant, particularly in patients with cochlear malformations, poorly pneumatized mastoid bones, or congenitally absent temporal bone landmarks, as well as those who have undergone previous ear surgery.

The lack of haptic feedback is a limitation of the *da Vinci Si* system. In standard master-slave–assisted laparoscopic procedures, surgeons rely on soft-tissue visual deformation to estimate the forces applied. Although the *da Vinci Si* workflow described in the present study is limited by the current lack of haptic feedback, this did not interfere with successful cochleostomy and mastoidectomy in both cases. Because the desired method of bone removal is to allow the high–revolutions per-minute drill bit (10,000–30,000 rpm) to progressively and smoothly mill and remove bone, the available visual feedback of the interaction of the drill bit at the bony surface was found in these two cadaveric specimens to be adequate for completion of these cases. It is our opinion that, because we did not need to rigorously press the drill bit burr against bone and because visual feedback of the system was deemed sufficient, the need for haptic feedback was mitigated. Additionally, we believe that the stereoscopic high-definition video and the audio feedback of the drill pitch helps to convey drill contact forces and interactions.

The fabrication of the 30° drill guide attachment is a key component of the designed system. Without this extension, there would be limited rotation of the robotic arm, which rotates the drill only in a single vertical axis thus decreases the degrees of freedom. Our 30° drill guide attachment to the robotic arm allows for rotation of the arm



and the drill to occur in a circular fashion along a vertical axis, directing a variable vector of the drill (dependent on the rotation). This markedly improved the ability of the drill to navigate around corners and allowed for completion of the surgical drilling in a manner very similar to freehand drilling.

More studies need to be performed to refine the surgical equipment for general widespread clinical use. We found three limitations of the technology that need to be addressed before considering taking this approach to clinical use. First, a smaller profile-articulating suction-irrigation device (Figure 6.14) for navigation in the facial recess is necessary. Similarly, instruments sufficiently small and specific for middle ear surgery need to be designed to operate with the robotic arm in the middle ear through the facial recess. At this point, microdissection in the middle ear with otologic instruments, including manipulation of the round window soft tissues, does not appear feasible given the access limitations. Second, improvement of the magnification of the 3D endoscope for improved visualization through the posterior tympanostomy is required because the view into the middle ear is better with classic binocular microscopy at this point. Third, once we have been able to complete this approach in more cadaveric specimens and have more adequately defined the accuracy, precision, and feasibility of the system, we plan to analyze the potential cost relative to the potential benefits of master-slave-assisted otologic surgery. Furthermore, though in this pilot feasibility study we did not measure the angle of insertion of the phantom array, future cadaveric studies should determine whether augmented reality can help predict the optimal angle and whether the position of the cochleostomy is adequate in all cases.

## 6.5 Chapter Summary and Future Work

We believe that both further technical and clinical assessments of our proposed system need to be conducted in order to evaluate its potential and overall effectiveness. Technical validation of our system includes assessment of individual component performance as well as overall usability studies of the entire system in realistic preclinical/clinical experiments. A critical question is whether the designed guidance system can potentially improve clinical utility. Several criteria are contemplated during the design of validation experiments including: (1) choosing appropriate data and phantom models to be used for particular surgical scenarios; (2) evaluating and validating systems and their individual components; and (3) testing different features of effective information delivery.

Generally, the lack of adequate validation and evaluation is a major obstacle to the clinical introduction of mixed paradigm image-guided robotic surgery (IGRS). Kersten-Oertel *et al.*<sup>298</sup> surveyed the current landscape and analyzed the chosen solutions for the components of related systems in several select publications. Results identify trends where a lack of focus in the assessment of these systems is common. Furthermore, the solutions presented are often based on available technology, anecdotes, and incomplete knowledge. System evaluations are performed in 87% of the selected publications, yet few papers looked at evaluating or validating the individual steps in the workflow and no publications examined all of the components. The majority of the researchers (36%) looked at validating the systems in terms of either the accuracy of the system as a whole, or the registration, calibration, or overlay accuracy of the real and virtual images. In general, although some system components were validated using numerical methods or

phantoms, the majority were not evaluated in clinical settings on real patients.

The specifications of these criteria and assessment levels outline the complexity and difficulty involved in assessing image-guided robotic surgical (IGRS) systems. Evaluating patient-related criteria in terms of surgical outcomes such as cosmetic results, pain, and clinical scores is particularly challenging and time consuming, for example, in planning and executing clinical trials then measuring improvement based on whether a particular technology or technique was used. New metrics for assessing the different criteria need to be developed. These require a focus on working with surgeons and physicians to define metrics that can be used not only to examine systems individually but also to compare their components and the systems themselves.

While standardizing technical and clinical benchmarks will facilitate the introduction of similar IGRS systems, historical success of prior related technology has heavily been dependent on regulatory approval (e.g., from the FDA, etc.) and patient demand, which is related to marketing strategies. In fact, Zender *et al.*<sup>299</sup> recommends that the first step in setting up a robotic surgery program at any facility, but particularly in a rural area, is to perform market research of the geographic area in order to assess whether there is a potential need. Arguably, flexibility in regulatory approval and patient demand has led to the success of the *da Vinci* system from its original intended design as a cardiothoracic system into the mainstream option for prostatectomies and hysterectomies<sup>300</sup>. There exist distinct advantages of the multi-market versatility of this platform, compared to single application robotic systems<sup>70,82,301</sup>. Therefore, we have conscientiously designed a modular architecture with interventional flexibility adaptable to a range of clinical applications, with this chapter focused on their individual

validations.

Our efforts have recognized that, at the onset of the development of a complex medical image guidance system, targeted user studies should be incrementally completed and aimed at learning the needs of surgeons and the constraints of the complex OR environment. Furthermore, efforts in validation should continuously re-assess both the performance limitations of each component and how they impact the overall system target accuracy. The variety of the publications reviewed in several surveys<sup>108,298,302</sup> demonstrates that many different technological solutions and tools exist for each of the individual components. Survey findings recommend that the focus in the field, must now turn from technical innovations to methods of combining these tools to develop systems that fit seamlessly into the OR and aid the surgeon in specific tasks. We have adopted such an approach in this dissertation and have worked closely with clinicians in order to determine the surgeons' requirements in the OR. We believe that the development of our system based on preclinical requirements, as well as the evaluation and validation of components and the overall system, as presented in this chapter and previous sections, will be essential for the next step in the successful introduction of such mixed paradigm IGRS systems into clinical consideration.

TABLE 6.4 RESEARCH PROBLEMS (SECTION 1.4) WITH TECHNICAL BARRIERS RESOLVED  
(SECTION 1.5) THROUGH INLINE CONTRIBUTIONS

- Problem 1: Mental Correspondence
  - Effective data representation from volumetric CT/MRI/CBCT
    - Augmented reality through video overlay of critical anatomies
  - Effective information delivery from tool positions and volumetric CT/MRI/CBCT
    - Augmented reality through enhanced stereoscopic depth perception
    - Augmented reality with tool tracking and localization with respect to critical anatomical information
  
- Problem 3: Effective Image Guidance
  - Effective systems engineering with evaluation and validation
    - Feasible and deployable clinical workflow
    - Evaluation of effectiveness and accuracy of individual features and methodology of image guidance
    - Pre-clinical validation with in vivo experiments for clinically-relevant scenarios

## 7 Summary and Conclusions

Currently, in minimally-invasive robotic surgery, a gap exists between incorporating medical *information* from standard preoperative diagnostic patient imaging data and intraoperative *action*. Clinicians must mentally map correspondence between dense volumetric CT/MRI to endoscopic findings, which is a highly subjective skill best practiced with normalized anatomy and based on extensive operative experience. Furthermore, when compared to open surgery, diminished visualization and reduced tactile feedback of elongated tools controlled through a fulcrum both create an environment that necessitates complex hand-eye coordination.

In an effort to address these fundamental clinical and technical limitations, The goal of our mixed paradigm computer-integrated surgical system is to provide image guidance based on interventional C-arm imaging, thereby adding detailed 3D information that is helpful for accurate navigation. Our design integrates patient-specific models derived from multi-modal preoperative medical data, intraoperative imaging with a high-end robotic C-arm, and the mechanical dexterity of current state-of-the-art robotic platforms (e.g., the *da Vinci* system). The approach involve the development of a versatile image-guided robotic surgical system that is modularly reconfigurable and the contribution of novel technical capabilities in augmented reality for guided robotic

surgery.

The Health Care Technology Assessment (HCTA) from the National Institutes of Health (NIH) states that "*Health technology is the practical application of knowledge.*" The HCTA cites a multi-level assessment process with three modes to describe health care technology, including its material nature, its purpose, and its stage of diffusion. In this dissertation, we have presented the basic level, describing the technical components contributing to the overall architecture. These include characterization of the intrinsic functionality and accuracy of each individual component, as well as the behavior of an entire system as a comprehensive image-guided robotic surgical system. The effectiveness of the design has been demonstrated in many *ex vivo* and *in vivo* operating scenarios. We believe that building a modular extendible system that can be applied across multiple surgical specialties and designed in close collaboration with clinicians will help ease distribution and diffusion of the proposed solutions to improve upon current clinical practice.

## **7.1 Future Improvements**

The ultimate goal of the work in this dissertation is to use the proposed image guidance system in mainstream surgery. Given the current state of development as presented, along with the future work discussed in each chapter, the following sections contain several future projects that need to be performed in order to further these efforts towards that goal.

### **7.1.1 Optimizing Modules and Functionalities**

Several modules and features of the IGRS system can be improved as follows:

#### **7.1.1.1 Augmented Reality**

In minimally invasive interventions, the guidance environment is the surgeon's only visual access to the surgical site, raising the following questions: What information is appropriate? How much information is sufficient? When and how should this information be displayed? How can the surgeon interact with the data? Augmented reality in image guidance for surgery therefore should inject information for users when needed; it should not diminish users' efficiency during the rest of the clinical workflow. Many studies have shown that visual perception with augmented reality, especially depth perception, is a psychophysical and heavily user-dependent experience based on human factors. Subtle graphical rendering techniques (e.g., shading, opacity, lighting, etc.) can significantly affect visual interpretation<sup>254,255</sup>. Future efforts should explore these techniques in order to develop optimal visualization and navigation paradigms through clinical multi-user studies.

#### **7.1.1.2 Improved Tool Tracking**

Navigation by tracking tools with respect to medical imaging is an active area of research<sup>267,303</sup>. Several prior studies<sup>173,304</sup> have noted a discrepancy in the accuracy of navigation systems in the laboratory versus a clinical setting. For example, EM tracking systems, required for instrument tracking inside the body, feature a localization accuracy of 2 mm or better<sup>305,306</sup> in certain clinical environments<sup>307</sup> with worse performances with EM interference<sup>175</sup>. Thus, these tool tracking solutions, including our hybrid kinematic and vision-based approach, must be thoroughly evaluated to show that they are usable not only in ideal circumstances, but also under clinically challenging conditions. Future work should look to consistently examine the effectiveness of our solutions, using the same



criteria to compare solutions across different systems.

### **7.1.1.3 Updates for Non-Rigid (Deformable) Tissue Deformation**

Understandably, application of image-guidance technology in general abdominal surgery (e.g., hepatic interventions) has been limited for various reasons, including motion deformation of abdominal soft and solid organs by insufflation, surgical manipulation, or ventilation. In abdominal surgery, because the target organ is likely continuously to move or be displaced in comparison with the time of preoperative image acquisition, augmented reality through video augmentation is considerably more challenging to achieve.

There has been much progress to date, but further work needs to address these issues. Other groups have shown that periodic motion, such as that of the heart<sup>62</sup> or lungs, can be modeled sufficiently and can update the 3D patient data, but the opportunity remains to model non-periodic changes, such as tissue resection, through vision-based methods. Future work should investigate methods to update navigational information after non-rigid tissue deformation from surgical motion. For TORS, one solution is the integration of information from forward modeling of the motions of the oral tongue<sup>308-310</sup>. In addition, real-time, high-fidelity, deformable registration<sup>153,311-313</sup> for medical imaging (as shown in a survey by Sotiras *et al.*<sup>314</sup>) remains an active and extensive area of research.

## **7.2 Future Research Directions**

A considerable amount of future work can significantly substantiate the value and versatility of this system beyond capabilities already explored in this dissertation. Topics,

such as multi-modality integration and closed loop IGRS, highlight some of the most difficult directions of research related to a systems approach to the introduction of technology in medicine.

The following sections contain potential research directions that we can envision at this juncture:

### **7.2.1 Multi-Modality Integration**

The deformable registration algorithm developed for TORS (Section 3.6.1) has been verified with cadaveric phantoms to appropriately align preoperative CT to perioperative CBCT<sup>153,266</sup>. However, by using an intensity-invariant hybrid approach, this implementation is primed for multi-modality registration. Registration of preoperative MRI to CBCT would be an interesting and compelling fusion of modalities to verify in future studies. Furthermore, real-time image acquisition (e.g., from intraoperative ultrasound, CT, or MRI) should be explored, depending on the goal of the target clinical application. Currently, the main drawbacks for this 3D-3D deformable registration algorithm are the initial manual segmentation and a performance time of several minutes. Both aspects must be addressed in future work.

### **7.2.2 Continuous Assessment in a Closed Loop System**

A closed-loop image guidance approach brings information back into the system in order to enforce and confirm performance, accuracy, and progress. Ideally, an IGRS such as the one proposed in this dissertation, would continuously self-regulate key functionalities after the initial setup. For example, regarding the registration of the surgical CAD/CAM with video, we have demonstrated appropriate perioperative

accuracy (i.e.,  $\leq 2$  mm). That said, surgical intervention progresses, and, while the fidelity of the overlaid anatomical models is expected to maintain similar reliability, our design does not currently include methods of continuously verifying accuracy. Other image-guided robotic systems, such as ROBODOC<sup>315,316</sup>, feature built-in redundant sensors and motion monitoring through optical tracking to ensure safety and consistency. Similar continuous assessment and feedback to the users can greatly improve surgical confidence for utilization of this system.

### **7.3 Summary**

This doctoral dissertation, titled “Augmented Reality and Intraoperative C-Arm Cone-Beam Computed Tomography for Image-Guided Robotic Surgery” explores new methods to apply augmented reality to integrate image guidance from intraoperative C-Arm-based imaging to assist surgeons in performing highly skilled minimally-invasive robotic surgery. We demonstrate the effectiveness of the proposed system for multiple clinical applications including transoral robotic surgery, robot-assisted thoracic surgery and robot-assisted mastoidectomy and cochleostomy. Transoral robotic surgery in particular, is a challenging discipline due to limited visual access of embedded oncologic targets, a lack of force feedback to palpate and delineate the tissue boundaries, and a gap in visualization between preoperative volumetric imaging (CT/MRI) and a possibly highly deformed surgical field.

The body of work implemented a modular architecture and the design principles for a C-Arm-guided robotic system that are used to develop innovative functionality to address these fundamental limitations in minimally-invasive robotic surgery. It is an inherently information driven modular system incorporating robotics, intraoperative imaging, and

multimedia components. This work integrates a high-end robotic C-Arm system and a modern surgical robotic system, whose integrated systems are leveraged and extended to create intuitive and relevant visualization, human-machine interferences in a streamlined approach. Guidance provided by this system uses augmented reality fusing virtual medical information, tool localization and other dynamic information behavior in order to present novel enhanced depth information to the surgeon. The resulting functionality, and the proposed architecture and design methods generalize to other C-Arm-based image guidance for robotic surgery. The system's performance is demonstrated and evaluated using phantoms and in-vivo experiments.

The main contributions reported in this work include:

1. A design and implementation of a component-based software architecture by extension of existing open source framework and libraries
2. The demonstration of modularity of the design in two distinctive architectural configurations, using changeable hardware interfaces, required for different clinical applications.
3. The novel systems integration of a robotic C-Arm system and multiple models of a surgical robotic platform.
4. Design and implementation of intraoperative updates for image-guided augmented reality using real-time imaging and vision-based techniques.
5. Novel technical methods in augmented reality for both monocular and stereoscopic video cameras that can be used to address specific surgical

challenges in visualization through mixed-reality derived from medical information using standard diagnostic volumetric CT/MRI/CBCT.

These include:

- a. Visual tool-to-critical structure proximity information feedback based on robot kinematics, and vision-based tracking of calibrated tools
  - b. Enhanced depth perception with novel orthogonal view with picture-in-picture inset into augmented stereo endoscopy
6. Design and creation of mock tumors and artificial ex vivo animal phantoms, cadaveric head phantoms and in vivo animal phantoms.
7. An experimental testbed for development, demonstration and evaluation of the overall system, and task-specific functionality.
8. Basic science experiments to assess methods of visualization using augmented reality for embedded target localization with stereoscopic robotic endoscopy.
9. Clinically-relevant, in vivo experiments to assess effectiveness of proposed methods of image guidance compared to simulated current practice.



# Bibliography

1. Abdullah A, Mahmud MR, Maimunah A, Zulfiqar MA, Saim L, Mazlan R. Preoperative high resolution CT and MR imaging in cochlear implantation. *Annals of the Academy of Medicine, Singapore*. Jul 2003;32(4):442-445.
2. Sahani D, Mehta A, Blake M, Prasad S, Harris G, Saini S. Preoperative hepatic vascular evaluation with CT and MR angiography: implications for surgery. *Radiographics*. Sep-Oct 2004;24(5):1367-1380.
3. Ozsarlak O, Tjalma W, Schepens E, et al. The correlation of preoperative CT, MR imaging, and clinical staging (FIGO) with histopathology findings in primary cervical carcinoma. *Eur Radiol*. Oct 2003;13(10):2338-2345.
4. Tanaka S, Matsuo K, Sasaki T, Nakano M, Shimura H, Yamashita Y. Clinical outcomes and advantages of laparoscopic surgery for primary Crohn's disease: are they significant? *Hepato-gastroenterology*. Mar-Apr 2009;56(90):416-420.
5. Schmelzer TM, Hope WW, Iannitti DA, Kercher KW, Heniford BT. Laparoscopic colostomy takedown offers advantages over traditional surgery. *Journal of minimal access surgery*. Dec 2006;2(4):201-202.
6. Feng B, Zheng MH, Mao ZH, et al. Clinical advantages of laparoscopic colorectal cancer surgery in the elderly. *Aging clinical and experimental research*. Jun 2006;18(3):191-195.
7. von Allmen D, Markowitz JE, York A, Mamula P, Shepanski M, Baldassano R. Laparoscopic-assisted bowel resection offers advantages over open surgery for treatment of segmental Crohn's disease in children. *Journal of pediatric surgery*. Jun 2003;38(6):963-965.
8. Yang CC, Chen P, Tseng JY, Wang PH. Advantages of open laparoscopic surgery over exploratory laparotomy in patients with tubo-ovarian abscess. *The Journal of the American Association of Gynecologic Laparoscopists*. Aug 2002;9(3):327-332.
9. Fried M, Peskova M, Kasalicky M, Svab J. Advantages of laparoscopic surgery in advanced pancreatic tumors. *Acta Universitatis Palackianae Olomucensis Facultatis Medicae*. 2000;143:82.
10. Diodato MD, Jr., Prosd SM, Klingensmith ME, Damiano RJ, Jr. Robotics in surgery. *Current problems in surgery*. Sep 2004;41(9):752-810.
11. Rank WP. Getting ahead with MRA and CTA. *Nursing*. Spring 2011;41 Suppl Insider:6-7.
12. Varma TR, Eldridge P. Use of the NeuroMate stereotactic robot in a frameless mode for functional neurosurgery. *The international journal of medical robotics + computer assisted surgery : MRCAS*. Jun 2006;2(2):107-113.
13. Rizun PR, McBeth PB, Louw DF, Sutherland GR. Robot-assisted neurosurgery. *Seminars in laparoscopic surgery*. Jun 2004;11(2):99-106.
14. Balicki M, Xia T, Jung MY, et al. Prototyping a Hybrid Cooperative and Tele-robotic Surgical System for Retinal Microsurgery. *MIDAS journal*. Jun 2011.
15. Barrett SF, Wright CH, Jerath MR, et al. Automated retinal robotic laser system. *Biomedical sciences instrumentation*. 1995;31:89-93.
16. Ortmaier T, Weiss H, Dobeles S, Schreiber U. Experiments on robot-assisted navigated drilling and milling of bones for pedicle screw placement. *The international journal of medical robotics + computer assisted surgery : MRCAS*. Dec 2006;2(4):350-363.
17. Zhao CP, Wang JQ, Liu WY, et al. [Accurate traction of long bone fracture with full-length planning module of orthopedic robot system: experiments in vitro and in vivo]. *Zhonghua yi xue za zhi*. Nov 20 2007;87(43):3038-3042.
18. Yamamura M, Nakamura N, Miki H, Nishii T, Sugano N. Cement Removal from the Femur Using the ROBODOC System in Revision Total Hip Arthroplasty. *Advances in orthopedics*. 2013;2013:347358.
19. Roche M. Robotic-assisted unicompartmental knee arthroplasty: the MAKO experience. *Clinics in sports medicine*. Jan 2014;33(1):123-132.
20. Ramqvist T, Dalianis T. Oropharyngeal cancer epidemic and human papillomavirus. *Emerging infectious diseases*. Nov 2010;16(11):1671-1677.
21. Isayeva T, Li Y, Maswahu D, Brandwein-Gensler M. Human papillomavirus in non-oropharyngeal head and neck cancers: a systematic literature review. *Head and neck pathology*. Jul 2012;6 Suppl 1:S104-120.



22. Westra WH. The changing face of head and neck cancer in the 21st century: the impact of HPV on the epidemiology and pathology of oral cancer. *Head and neck pathology*. Mar 2009;3(1):78-81.
23. Zhen W, Karnell LH, Hoffman HT, Funk GF, Buatti JM, Menck HR. The National Cancer Data Base report on squamous cell carcinoma of the base of tongue. *Head Neck*. Aug 2004;26(8):660-674.
24. Weinstein GS, O'Malley BW, Jr., Magnuson JS, et al. Transoral robotic surgery: a multicenter study to assess feasibility, safety, and surgical margins. *Laryngoscope*. Aug 2012;122(8):1701-1707.
25. Weinstein GS, Quon H, Newman HJ, et al. Transoral robotic surgery alone for oropharyngeal cancer: an analysis of local control. *Archives of otolaryngology--head & neck surgery*. Jul 2012;138(7):628-634.
26. U.K. CR. <http://www.cancerresearchuk.org/>. 2013. Accessed 4 June, 2014.
27. Farivar AS, Cerfolio RJ, Vallieres E, et al. Comparing robotic lung resection with thoracotomy and video-assisted thoracoscopic surgery cases entered into the society of thoracic surgeons database. *Innovations (Phila)*. Jan-Feb 2014;9(1):10-15.
28. Kent M, Wang T, Whyte R, Curran T, Flores R, Gangadharan S. Open, video-assisted thoracic surgery, and robotic lobectomy: review of a national database. *Ann Thorac Surg*. Jan 2014;97(1):236-244.
29. Park BJ. Robotic lobectomy for non-small cell lung cancer (NSCLC): Multi-center registry study of long-term oncologic results. *Annals of cardiothoracic surgery*. May 2012;1(1):24-26.
30. Pardolesi A, Park B, Petrella F, Borri A, Gasparri R, Veronesi G. Robotic anatomic segmentectomy of the lung: technical aspects and initial results. *Ann Thorac Surg*. Sep 2012;94(3):929-934.
31. Swanson SJ. Video-assisted thoracic surgery segmentectomy: the future of surgery for lung cancer? *Ann Thorac Surg*. Jun 2010;89(6):S2096-2097.
32. Snyderman CH, Pant H, Carrau RL, Prevedello D, Gardner P, Kassam AB. What are the limits of endoscopic sinus surgery?: the expanded endonasal approach to the skull base. *The Keio journal of medicine*. Sep 2009;58(3):152-160.
33. Magazine MT. <http://www.medicaltourismmag.com/>. 2014. Accessed 4 June, 2014.
34. Niparko JK. The significance of cochlear implant history. *JAMA otolaryngology-- head & neck surgery*. May 2013;139(5):454.
35. Skinner MW, Holden TA, Whiting BR, et al. In vivo estimates of the position of advanced bionics electrode arrays in the human cochlea. *The Annals of otology, rhinology & laryngology. Supplement*. Apr 2007;197:2-24.
36. Briggs RJ, Tykocinski M, Stidham K, Roberson JB. Cochleostomy site: implications for electrode placement and hearing preservation. *Acta oto-laryngologica*. Aug 2005;125(8):870-876.
37. Aschendorff A, Kromeier J, Klenzner T, Laszig R. Quality control after insertion of the nucleus contour and contour advance electrode in adults. *Ear and hearing*. Apr 2007;28(2 Suppl):75S-79S.
38. Adunka OF, Buchman CA. Scala tympani cochleostomy I: results of a survey. *Laryngoscope*. Dec 2007;117(12):2187-2194.
39. Adunka OF, Radeloff A, Gstoettner WK, Pillsbury HC, Buchman CA. Scala tympani cochleostomy II: topography and histology. *Laryngoscope*. Dec 2007;117(12):2195-2200.
40. Zhang J, Wei W, Manolidis S, Roland JT, Jr., Simaan N. Path planning and workspace determination for robot-assisted insertion of steerable electrode arrays for cochlear implant surgery. *Med Image Comput Comput Assist Interv*. 2008;11(Pt 2):692-700.
41. Lim H, J. H, Hong J, et al. Image-guided robotic mastoidectomy using human-robot collaboration control. International Conference on Mechatronics and Automation (ICMA); 2011; Beijing, China.
42. Becker BC, Maclachlan RA, Hager GD, Riviere CN. Handheld Micromanipulation with Vision-Based Virtual Fixtures. *IEEE International Conference on Robotics and Automation : ICRA : [proceedings] IEEE International Conference on Robotics and Automation*. Aug 18 2011;2011:4127-4132.

43. Zheng B, Kuang A, Henigman F, et al. Effects of assembling virtual fixtures on learning a navigation task. *Studies in health technology and informatics*. 2006;119:586-591.
44. Ren J, Patel RV, McIsaac KA. Rendering of virtual fixtures for MIS using generalized sigmoid functions. *Studies in health technology and informatics*. 2006;119:446-448.
45. Marayong P, Okamura AM. Speed-accuracy characteristics of human-machine cooperative manipulation using virtual fixtures with variable admittance. *Human factors*. Fall 2004;46(3):518-532.
46. KidsHealth. <http://kidshealth.org/>. 2014. Accessed 4 June, 2014.
47. Volonte F, Buchs NC, Pugin F, et al. Stereoscopic augmented reality for da Vinci robotic biliary surgery. *International journal of surgery case reports*. 2013;4(4):365-367.
48. Volonte F, Buchs NC, Pugin F, et al. Augmented reality to the rescue of the minimally invasive surgeon. The usefulness of the interposition of stereoscopic images in the Da Vinci robotic console. *The international journal of medical robotics + computer assisted surgery : MRCAS*. Sep 2013;9(3):e34-38.
49. Su LM, Vagvolgyi BP, Agarwal R, Reiley CE, Taylor RH, Hager GD. Augmented reality during robot-assisted laparoscopic partial nephrectomy: toward real-time 3D-CT to stereoscopic video registration. *Urology*. Apr 2009;73(4):896-900.
50. Hughes-Hallett A, Mayer EK, Marcus HJ, et al. Augmented Reality Partial Nephrectomy: Examining the Current Status and Future Perspectives. *Urology*. Oct 19 2013.
51. Badani KK, Shapiro EY, Berg WT, et al. A Pilot Study of Laparoscopic Doppler Ultrasound Probe to Map Arterial Vascular Flow within the Neurovascular Bundle during Robot-Assisted Radical Prostatectomy. *Prostate cancer*. 2013;2013:810715.
52. Chen X, Wang L, Fallavollita P, Navab N. Precise X-ray and video overlay for augmented reality fluoroscopy. *Int J Comput Assist Radiol Surg*. Jan 2013;8(1):29-38.
53. Van de Kelft E, Costa F, Van der Planken D, Schils F. A prospective multicenter registry on the accuracy of pedicle screw placement in the thoracic, lumbar, and sacral levels with the use of the O-arm imaging system and StealthStation Navigation. *Spine*. Dec 1 2012;37(25):E1580-1587.
54. Mirota DJ, Taylor RH, Ishii M, Hager GD. Direct endoscopic video registration for sinus surgery. SPIE Medical Imaging; 2009; San Diego, CA.
55. Cobb J, Henckel J, Gomes P, et al. Hands-on robotic unicompartmental knee replacement: a prospective, randomised controlled study of the acrobot system. *The Journal of bone and joint surgery. British volume*. Feb 2006;88(2):188-197.
56. Uneri A, Balicki MA, Handa J, Gehlbach P, Taylor RH, Iordachita I. New Steady-Hand Eye Robot with Micro-Force Sensing for Vitreoretinal Surgery. *Proceedings of the ... IEEE/RAS-EMBS International Conference on Biomedical Robotics and Biomechatronics. IEEE/RAS-EMBS International Conference on Biomedical Robotics and Biomechatronics*. Sep 1 2010;2010(26-29):814-819.
57. Li QH, Zamorano L, Pandya A, Perez R, Gong J, Diaz F. The application accuracy of the NeuroMate robot--A quantitative comparison with frameless and frame-based surgical localization systems. *Comput Aided Surg*. 2002;7(2):90-98.
58. Kazanzides P, Fichtinger G, Hager GD, Okamura AM, Whitcomb LL, Taylor RH. Surgical and Interventional Robotics: Core Concepts, Technology, and Design. *IEEE robotics & automation magazine / IEEE Robotics & Automation Society*. Jun 1 2008;15(2):122-130.
59. Taylor RH, Stoianovici D. Medical Robotics in Computer-Integrated Surgery. *IEEE Trans On Robotics and Automation*. 2003;19(5).
60. Freschi C, Troia E, Ferrari V, Megali G, Pietrabissa A, Mosca F. Ultrasound guided robotic biopsy using augmented reality and human-robot cooperative control. *Conf Proc IEEE Eng Med Biol Soc*. 2009;2009:5110-5113.
61. Pietrabissa A, Morelli L, Ferrari M, et al. Mixed reality for robotic treatment of a splenic artery aneurysm. *Surg Endosc*. May 2010;24(5):1204.
62. Figl M, Rueckert D, Hawkes D, et al. Image guidance for robotic minimally invasive coronary artery bypass. *Comput Med Imaging Graph*. Jan 2010;34(1):61-68.
63. Lanfranco AR, Castellanos AE, Desai JP, Meyers WC. Robotic surgery: a current perspective. *Ann Surg*. Jan 2004;239(1):14-21.

64. Kwoh YS, Hou J, Jonckheere EA, Hayati S. A robot with improved absolute positioning accuracy for CT guided stereotactic brain surgery. *IEEE Trans Biomed Eng.* Feb 1988;35(2):153-160.
65. Davies B. A review of robotics in surgery. *Proceedings of the Institution of Mechanical Engineers. Part H, Journal of engineering in medicine.* 2000;214(1):129-140.
66. Taylor RH, Joskowicz L, Williamson B, et al. Computer-integrated revision total hip replacement surgery: concept and preliminary results. *Med Image Anal.* Sep 1999;3(3):301-319.
67. Spencer EH. The ROBODOC clinical trial: a robotic assistant for total hip arthroplasty. *Orthopaedic nursing / National Association of Orthopaedic Nurses.* Jan-Feb 1996;15(1):9-14.
68. Cowley G. Introducing "Robodoc". A robot finds his calling--in the operating room. *Newsweek.* Nov 23 1992;120(21):86.
69. Song EK, Seon JK, Yim JH, Netravali NA, Bargar WL. Robotic-assisted TKA reduces postoperative alignment outliers and improves gap balance compared to conventional TKA. *Clin Orthop Relat Res.* Jan 2013;471(1):118-126.
70. Werner SD, Stonestreet M, Jacofsky DJ. Makoplasty and the Accuracy and Efficacy of Robotic-assisted Arthroplasty. *Surgical technology international.* Mar 2014;24:302-306.
71. Pearle AD, Kendoff D, Stueber V, Musahl V, Repicci JA. Perioperative management of unicompartmental knee arthroplasty using the MAKO robotic arm system (MAKOplasty). *American journal of orthopedics.* Feb 2009;38(2 Suppl):16-19.
72. Jaramaz B, Picard F, Gregori A. EVALUATING JOINT LAXITY IN UKR USING NAVIO-PFS. Paper presented at: Orthopaedic Proceedings, A supplement to The Bone & Joint Journal 2013.
73. Cheng L, Cao R, Meng G, Huang Q, Hou D, Hu L. Application of computer assisted navigation system in endoscopic sinus and skull base surgery. *Lin Chung Er Bi Yan Hou Tou Jing Wai Ke Za Zhi.* Sep 2012;26(17):796-798.
74. Gumprecht HK, Widenka DC, Lumenta CB. BrainLab VectorVision Neuronavigation System: technology and clinical experiences in 131 cases. *Neurosurgery.* Jan 1999;44(1):97-104; discussion 104-105.
75. Marichal JF, Devries C, Aime F, et al. [Decentralization of the ARTEMIS system in a nephrology service in a general hospital]. *Archives des maladies du coeur et des vaisseaux.* Jun 1987;80(6):888-891.
76. NASA technology finds earthly use: robot will probe brain for tumors. *Biomedical instrumentation & technology / Association for the Advancement of Medical Instrumentation.* Mar-Apr 1997;31(2):170.
77. Pavic R. Use of the trauma pelvic orthotic device (T-POD) for provisional stabilisation of anterior-posterior compression type pelvic fractures: a cadaveric study [Injury, Int J Care Injured 2008;39:903-6]. *Injury.* Jun 2009;40(6):673; author reply 673-674.
78. Garcia P, Rosen J, Kapoor C, et al. Trauma Pod: a semi-automated telerobotic surgical system. *The international journal of medical robotics + computer assisted surgery : MRCAS.* Jun 2009;5(2):136-146.
79. Satava RM. Surgical robotics: the early chronicles: a personal historical perspective. *Surgical laparoscopy, endoscopy & percutaneous techniques.* Feb 2002;12(1):6-16.
80. Taylor RH, Funda J, Eldridge B, et al. A telerobotic assistant for laparoscopic surgery. *IEEE engineering in medicine and biology magazine : the quarterly magazine of the Engineering in Medicine & Biology Society.* 1995:279-291.
81. Schreiner S, Anderson J, Taylor R, Funda J, Bzostek A, Barnes A. A system for percutaneous delivery of treatment with a fluoroscopically-guided robot. Paper presented at: Joint Conf. Computer Vision, Virtual Reality, and Robotics in Medicine and Medical Robotics and Computer Surgery 1997; Grenoble, France.
82. Pugin F, Bucher P, Morel P. History of robotic surgery: from AESOP(R) and ZEUS(R) to da Vinci(R). *Journal of visceral surgery.* Oct 2011;148(5 Suppl):e3-8.
83. Reichenspurner H, Damiano RJ, Mack M, et al. Use of the voice-controlled and computer-assisted surgical system ZEUS for endoscopic coronary artery bypass grafting. *J Thorac Cardiovasc Surg.* Jul 1999;118(1):11-16.

84. Marescaux J, Leroy J, Rubino F, et al. Transcontinental robot-assisted remote telesurgery: feasibility and potential applications. *Ann Surg.* Apr 2002;235(4):487-492.
85. Satava RM, Simon IB. Teleoperation, telerobotics, and telepresence in surgery. *Endoscopic surgery and allied technologies.* Jun 1993;1(3):151-153.
86. Guthart GS, Salisbury JK. The intuitive telesurgery system: Overview and application. Paper presented at: IEEE International Conference Robotics and Automation2000; San Francisco.
87. Poliachenko Iu V, Nychytailo M, Kostyliov MV, Zahriichuk MS. Robotic-assisted mini-invasive surgery: history, modern state, perspectives. *Klinichna khirurgiia / Ministerstvo okhorony zdorov'ia Ukrainy, Naukove tovarystvo khirurhiv Ukrainy.* Apr 2011(4):60-64.
88. Allendorf JD, Bessler M, Whelan RL, et al. Postoperative immune function varies inversely with the degree of surgical trauma in a murine model. *Surg Endosc.* May 1997;11(5):427-430.
89. Martino MA, Berger EA, McFetridge JT, et al. A Comparison of Quality Outcome Measures in Patients Having a Hysterectomy for Benign Disease: Robotic vs. Non-robotic Approaches. *The Journal of Minimally Invasive Gynecology.* 2013.
90. MA P, BB M, UK J. National multi-institutional comparison of 30-day post-operative complication and re-admission rates between open retropubic radical prostatectomy (RRP) and robot-assisted laparoscopic prostatectomy (RALRP) using NSQIP. *J Endourology.* 2013.
91. Faria EF, Caputo PA, Wood CG, Karam JA, Noguera-Gonzalez GM, Matin SF. Robotic partial nephrectomy shortens warm ischemia time, reducing suturing time kinetics even for an experienced laparoscopic surgeon: a comparative analysis. *World journal of urology.* Feb 2014;32(1):265-271.
92. Adelstein DJ, Ridge JA, Brizel DM, et al. Transoral resection of pharyngeal cancer: summary of a National Cancer Institute Head and Neck Cancer Steering Committee Clinical Trials Planning Meeting, November 6-7, 2011, Arlington, Virginia. *Head Neck.* Dec 2012;34(12):1681-1703.
93. Van Abel KM, Moore EJ. The rise of transoral robotic surgery in the head and neck: emerging applications. *Expert review of anticancer therapy.* Mar 2012;12(3):373-380.
94. Nicolau S, Soler L, Mutter D, Marescaux J. Augmented reality in laparoscopic surgical oncology. *Surgical oncology.* Sep 2011;20(3):189-201.
95. Ukimura O, Gill IS. Image-fusion, augmented reality, and predictive surgical navigation. *The Urologic clinics of North America.* May 2009;36(2):115-123, vii.
96. Mirota D, Wang H, Taylor RH, Ishii M, Hager GD. Toward video-based navigation for endoscopic endonasal skull base surgery. *Med Image Comput Comput Assist Interv.* 2009;12(Pt 1):91-99.
97. Mirota DJ, Uneri A, Schafer S, et al. High-accuracy 3D image-based registration of endoscopic video to C-arm cone-beam CT for image-guided skull base surgery. *SPIE Medical Imaging;* 2011; Lake Buena Vista, FL.
98. Liu WP, Mirota DJ, Uneri A, et al. A clinical pilot study of a modular video-CT augmentation system for image-guided skull base surgery. Paper presented at: SPIE Medical Imaging 2012: Image-Guided Procedures, Robotic Interventions, and Modeling2012.
99. Feuerstein M, Mussack T, Heining SM, Navab N. Intraoperative laparoscope augmentation for port placement and resection planning in minimally invasive liver resection. *IEEE Trans Med Imaging.* Mar 2008;27(3):355-369.
100. Long JA, Lee BH, Guillotreau J, et al. Real-time robotic transrectal ultrasound navigation during robotic radical prostatectomy: initial clinical experience. *Urology.* Sep 2012;80(3):608-613.
101. Han M, Kim C, Mozer P, et al. Tandem-robot assisted laparoscopic radical prostatectomy to improve the neurovascular bundle visualization: a feasibility study. *Urology.* Feb 2011;77(2):502-506.
102. Herrell SD, Kwartowitz DM, Milhoua PM, Galloway RL. Toward image guided robotic surgery: system validation. *J Urol.* Feb 2009;181(2):783-789; discussion 789-790.

103. Teber D, Guven S, Simpfendorfer T, et al. Augmented reality: a new tool to improve surgical accuracy during laparoscopic partial nephrectomy? Preliminary in vitro and in vivo results. *Eur Urol.* Aug 2009;56(2):332-338.
104. Fuchs H, Livingston MA, Raskar R, et al. Augmented Reality Visualization for Laparoscopic Surgery. *Medical Image Computing and Computer-Assisted Intervention*; 1998.
105. Volonte F, Pugin F, Bucher P, Sugimoto M, Ratib O, Morel P. Augmented reality and image overlay navigation with OsiriX in laparoscopic and robotic surgery: not only a matter of fashion. *Journal of hepato-biliary-pancreatic sciences.* Jul 2011;18(4):506-509.
106. Volonte F, Buchs NC, Pugin F, et al. Stereoscopic augmented reality for da Vinci robotic biliary surgery. *International journal of surgery case reports.* 2013;4(4):365-367.
107. Ukimura O, Magi-Galluzzi C, Gill IS. Real-time transrectal ultrasound guidance during laparoscopic radical prostatectomy: impact on surgical margins. *J Urol.* Apr 2006;175(4):1304-1310.
108. Linte CA, Davenport KP, Cleary K, et al. On mixed reality environments for minimally invasive therapy guidance: systems architecture, successes and challenges in their implementation from laboratory to clinic. *Comput Med Imaging Graph.* Mar 2013;37(2):83-97.
109. Suzuki N, Hattori A, Hashizume M. Benefits of augmented reality function for laparoscopic and endoscopic surgical robot systems. Paper presented at: Augmented environments for Medical Imaging including Augmented Reality in Computer-aided Surgery2008.
110. Hattori A, Suzuki N. A Robotic Surgery System (daVinci) with a Data Fusion System for Navigation Surgery in the Abdominal Region. *Jikeikai Medical Journal.* 2003;50:19-27.
111. Nithiananthan S, Brock KK, Daly MJ, Chan H, Irish JC, Siewerdsen JH. Demons deformable registration for CBCT-guided procedures in the head and neck: convergence and accuracy. *Med Phys.* Oct 2009;36(10):4755-4764.
112. Schafer S, Nithananiathan S, Mirota DJ, et al. Mobile C-Arm Cone-Beam CT for Guidance of Spine Surgery: Image Quality, Radiation Dose, and Integration with Interventional Guidance. *Medical Physics.* 2011;38(8):4563-4575.
113. Schafer S, Uneri A, Mirota DJ, et al. C-arm cone-beam CT guidance of video-assisted thoracoscopic surgery (VATS): Image quality, dose, and integration with interventional guidance. *ISCAS*; 2011, 2011; Berlin, Germany.
114. Siewerdsen JH, Daly MJ, Bachar G, et al. Multimode C-arm fluoroscopy, tomosynthesis, and cone-beam CT for image-guided interventions: from proof of principle to patient protocols. *SPIE Medical Imaging* 2007.
115. Eskander A, King E, Goldstein DP, Siewerdsen JH, Irish JC. Cone-Beam CT for image guided surgery of the head and neck. *University of Toronto Medical Journal.* 2008 2008;86(1):21-24.
116. Siewerdsen JH, Daly MJ, Chan H, Nithananiathan S, Hamming NM, Irish JC. Development and implementation of an integrated image-guided surgery system based on a prototype C-arm for high-performance cone-beam CT. *ISCAS*; 2009, 2009; Berlin, Germany.
117. Siewerdsen JH, Daly MJ, Chan H, et al. High-performance intraoperative cone-beam CT on a mobile C-arm: an integrated system for guidance of head and neck surgery. *SPIE Medical Imaging* 2009.
118. Enquobahrie A, Kelly M, Deguet A, et al. Robot assisted prostate surgery using augmented reality with deformable models. Paper presented at: MICCAI Workshop on Systems and Arch. for Computer Assisted Interventions2012.
119. Bethea BT, Okamura AM, Kitagawa M, et al. Application of haptic feedback to robotic surgery. *Journal of laparoendoscopic & advanced surgical techniques. Part A.* Jun 2004;14(3):191-195.
120. Fitzpatrick JM, West JB, Maurer CR, Jr. Predicting error in rigid-body point-based registration. *IEEE Trans Med Imaging.* Oct 1998;17(5):694-702.
121. Nason RW, Binahmed A, Pathak KA, Abdoh AA, Sandor GK. What is the adequate margin of surgical resection in oral cancer? *Oral Surg Oral Med Oral Pathol Oral Radiol Endod.* May 2009;107(5):625-629.

122. Elfring R, de la Fuente M, Radermacher K. Assessment of optical localizer accuracy for computer aided surgery systems. *Comput Aided Surg.* 2010;15(1-3):1-12.
123. Yaniv Z, Wilson E, Lindisch D, Cleary K. Electromagnetic tracking in the clinical environment. *Med Phys.* Mar 2009;36(3):876-892.
124. Lum M, Rosen J, Lendvay T, Sinanan M, Hannaford B. Effect of time delay on surgical performance during telesurgical manipulation. *IEEE International Conference on Robotics and Automation.*; 2009.
125. Balicki M. *Augmentation Of Human Skill In Microsurgery.* Baltimore, MD: Computer Science, Johns Hopkins University; 2014.
126. Uneri A, Schafer S, Mirota DJ, et al. TREK: an integrated system architecture for intraoperative cone-beam CT-guided surgery. *Int J Comput Assist Radiol Surg.* Jan 2012;7(1):159-173.
127. Sutherland I. A head-mounted three dimensional display. Paper presented at: Proceedings of the Fall Joint Compute Conference 1968.
128. Muensterer OJ, Lacher M, Zoeller C, Bronstein M, Kubler J. Google Glass in pediatric surgery: An exploratory study. *International journal of surgery.* 2014;12(4):281-289.
129. Deguet A, Kumar R, Taylor RH, Kazanzides P. The cisst libraries for computer assisted intervention systems. *MICCAI Workshop.* 2008; <https://trac.lcsr.jhu.edu/cisst/>.
130. Bowman DA, Coquillart S, Froehlich B, et al. 3D user interfaces: new directions and perspectives. *IEEE computer graphics and applications.* Nov-Dec 2008;28(6):20-36.
131. Traub J, Stefan P, Heining SM, et al. Hybrid navigation interface for orthopedic and trauma surgery. *Med Image Comput Comput Assist Interv.* 2006;9(Pt 1):373-380.
132. Pieper S, Lorenson B, Schroeder W, Kikinis R. The NA-MIC kit: ITK, VTK, pipelines, grids, and 3D Slicer as an open platform for the medical image computing community. *Proc. IEEE Intl. Symp. Biomed. Imag.* 2006(April):698-701.
133. Kwartowitz DM, Miga MI, Herrell SD, Galloway RL. Towards image guided robotic surgery: multi-arm tracking through hybrid localization. *Int J Comput Assist Radiol Surg.* May 2009;4(3):281-286.
134. Brennecke T, Jansen N, Raczekowski J, Schipper J, Woern H. An ultrasound-based navigation system for minimally invasive neck surgery. *Stud Health Technol Inform.* 2014;196:36-42.
135. Yang L, Wang J, Kobayashi E, et al. Ultrasound image-based endoscope localization for minimally invasive fetoscopic surgery. *Conf Proc IEEE Eng Med Biol Soc.* 2013;2013:1410-1413.
136. Kedziorek DA, Solaiyappan M, Walczak P, et al. Using C-arm x-ray imaging to guide local reporter probe delivery for tracking stem cell engraftment. *Theranostics.* 2013;3(11):916-926.
137. Yoo J, Schafer S, Uneri A, Otake Y, Khanna AJ, Siewerdsen JH. An electromagnetic "Tracker-in-Table" configuration for X-ray fluoroscopy and cone-beam CT-guided surgery. *Int J Comput Assist Radiol Surg.* May 15 2012.
138. Delpon G, Llagostera C, Le Blanc M, et al. [Use of IGRT for prostate cancers (OBI-CBCT Varian, ExacTrac BrainLAB and MVCT TomoTherapy)]. *Cancer radiotherapie : journal de la Societe francaise de radiotherapie oncologique.* Sep 2009;13(5):399-407.
139. Schouten R, Lee R, Boyd M, et al. Intra-operative cone-beam CT (O-arm) and stereotactic navigation in acute spinal trauma surgery. *Journal of clinical neuroscience : official journal of the Neurosurgical Society of Australasia.* Aug 2012;19(8):1137-1143.
140. Miracle AC, Mukherji SK. Conebeam CT of the Head and Neck, Part 2: Clinical Applications. *Am J Neuroradiol.* Aug 2009;30(7):1285-1292.
141. Keen HI, Mease PJ, Bingham CO, 3rd, Giles JT, Kaeley G, Conaghan PG. Systematic review of MRI, ultrasound, and scintigraphy as outcome measures for structural pathology in interventional therapeutic studies of knee arthritis: focus on responsiveness. *The Journal of rheumatology.* Jan 2011;38(1):142-154.
142. Song SE, Cho NB, Fischer G, et al. Development of a Pneumatic Robot for MRI-guided Transperineal Prostate Biopsy and Brachytherapy: New Approaches. *IEEE International Conference on Robotics and Automation : ICRA : [proceedings] IEEE International Conference on Robotics and Automation.* Jul 15 2010;2010:2580-2585.

143. Ricke J, Thormann M, Ludewig M, et al. MR-guided liver tumor ablation employing open high-field 1.0T MRI for image-guided brachytherapy. *Eur Radiol.* Aug 2010;20(8):1985-1993.
144. Siemens' Artis zeego brings surgery and industry together. *Cardiovasc J Afr.* Jul-Aug 2009;20(4):258.
145. Sadek M, Berland TL, Maldonado TS, et al. Use of preoperative magnetic resonance angiography and the Artis zeego fusion program to minimize contrast during endovascular repair of an iliac artery aneurysm. *Annals of vascular surgery.* Jan 2014;28(1):261 e261-265.
146. Strobl KH, Sepp W, Fuchs S, Paredes C, Arbter K. DLR CalDe and DLR CalLab. <http://www.robotic.dlr.de/callab/>. 2010.
147. Bradski G. The OpenCV Library. *Dr. Dobb's Journal of Software Tools.* 2000.
148. Daniilidis K. Hand-Eye Calibration Using Dual Quaternions. *The International journal of robotics research.* 1999;18(3):286-298.
149. Fahrig R, Moreau M, Holdsworth DW. Three-dimensional computed tomographic reconstruction using a C-arm mounted XRII: correction of image intensifier distortion. *Med Phys.* Jul 1997;24(7):1097-1106.
150. Livyatan H, Yaniv Z, L. J. Robust automatic c-arm calibration for fluoroscopy-based navigation: A practical approach. Paper presented at: MICCAI2002.
151. Yao J, Taylor RH, Goldberg RP, et al. Robust automatic c-arm calibration for fluoroscopybased navigation: A practical approach. *Computer Aided Surgery.* 2000;5(6):373-390.
152. Chintalapani G, Taylor RH. C-ARM DISTORTION CORRECTION USING PATIENT CT AS A FIDUCIAL. Paper presented at: IEEE International Symposium on Biomedical Imaging: From Nano to Macro2007.
153. Reaungamornrat S, Liu WP, Wang AS, et al. Deformable image registration for cone-beam CT guided transoral robotic base-of-tongue surgery. *Phys Med Biol.* Jul 21 2013;58(14):4951-4979.
154. Jian B, Vemuri BC. Robust Point Set Registration Using Gaussian Mixture Models. *IEEE Pattern Analysis and Machine Intelligence.* 2011;33(8):1633-1645.
155. Maurer CR, Jr., Rensheng Q, Raghavan V. A linear time algorithm for computing exact Euclidean distance transforms of binary images in arbitrary dimensions. *IEEE Pattern Analysis and Machine Intelligence.* 2003;25(2):265-270.
156. Vercauteren T, Pennec X, Malis E, Perchant A, Ayache N. Insight into Efficient Image Registration Techniques and the Demons Algorithm. *Proc. of Inf Process Med Imaging.* 2007;4584:495-506.
157. Reaungamornrat S, Liu WP, Schafer S, et al. A Gaussian Mixture + Demons Deformable Registration Method for Cone-Beam CT-Guided Robotic Transoral Base-of-Tongue Surgery. Paper presented at: SPIE Medical Imaging2013; Orlando.
158. Vescan AD, Chan H, Daly MJ, Witterick I, Irish JC, Siewerdsen JH. C-arm cone beam CT guidance of sinus and skull base surgery: quantitative surgical performance evaluation and development of a novel high-fidelity phantom. 2009;7261:72610L-72610.
159. Baumhauer M, Feuerstein M, Meinzer HP, Rassweiler J. Navigation in endoscopic soft tissue surgery: perspectives and limitations. *J Endourol.* Apr 2008;22(4):751-766.
160. Acosta FL, Jr., Thompson TL, Campbell S, Weinstein PR, Ames CP. Use of intraoperative isocentric C-arm 3D fluoroscopy for sextant percutaneous pedicle screw placement: case report and review of the literature. *Spine J.* May-Jun 2005;5(3):339-343.
161. Yao J, Taylor RH, Goldberg RP, et al. A C-arm fluoroscopy-guided progressive cut refinement strategy using a surgical robot. *Comput Aided Surg.* 2000;5(6):373-390.
162. Erovic BM, Daly MJ, Chan HH, et al. Evaluation of intraoperative cone beam computed tomography and optical drill tracking in temporal bone surgery. *Laryngoscope.* Nov 2013;123(11):2823-2828.
163. Yoshida G, Kanemura T, Ishikawa Y. Percutaneous Pedicle Screw Fixation of a Hangman's Fracture Using Intraoperative, Full Rotation, Three-dimensional Image (O-arm)-based Navigation: A Technical Case Report. *Asian spine journal.* Sep 2012;6(3):194-198.
164. Friedman WH, Katsantonis GP. Intranasal and transantral ethmoidectomy: a 20-year experience. *Laryngoscope.* Apr 1990;100(4):343-348.



165. Moore CE, Ross DA, Marentette LJ. Subcranial approach to tumors of the anterior cranial base: analysis of current and traditional surgical techniques. *Otolaryngol Head Neck Surg.* Mar 1999;120(3):387-390.
166. King AP, Edwards PJ, Maurer CR, Jr., et al. A system for microscope-assisted guided interventions. *Stereotactic and functional neurosurgery.* 1999;72(2-4):107-111.
167. Anon JB, Klimek L, Mosges R, Zinreich SJ. Computer-assisted endoscopic sinus surgery. An international review. *Otolaryngologic clinics of North America.* Jun 1997;30(3):389-401.
168. Stelter K, Andratschke M, Leunig A, Hagedorn H. Computer-assisted surgery of the paranasal sinuses: technical and clinical experience with 368 patients, using the Vector Vision Compact system. *The Journal of laryngology and otology.* Dec 2006;120(12):1026-1032.
169. Hauser R, Westermann B, Reinhardt H, Probst R. [Computer-assisted surgery of the paranasal sinuses with an opto-electronic stereotaxic system]. *Laryngo- rhino- otologie.* Apr 1996;75(4):199-207.
170. Mosges R, Klimek L. Computer-assisted surgery of the paranasal sinuses. *The Journal of otolaryngology.* Apr 1993;22(2):69-71.
171. Fried MP, Kleefield J, Gopal H, Reardon E, Ho BT, Kuhn FA. Image-guided endoscopic surgery: results of accuracy and performance in a multicenter clinical study using an electromagnetic tracking system. *Laryngoscope.* May 1997;107(5):594-601.
172. Maciunas RJ, Galloway RL, Jr., Fitzpatrick JM, Mandava VR, Edwards CA, Allen GS. A universal system for interactive image-directed neurosurgery. *Stereotactic and functional neurosurgery.* 1992;58(1-4):108-113.
173. Ecke U, Luebben B, Maurer J, Boor S, Mann WJ. Comparison of Different Computer-Aided Surgery Systems in Skull Base Surgery. *Skull base : official journal of North American Skull Base Society ... [et al.].* Feb 2003;13(1):43-50.
174. Glossop ND. Advantages of optical compared with electromagnetic tracking. *J Bone Joint Surg Am.* Feb 2009;91 Suppl 1:23-28.
175. Reichl T, Gardiazabal J, Navab N. Electromagnetic servoing-a new tracking paradigm. *IEEE Trans Med Imaging.* Aug 2013;32(8):1526-1535.
176. Hauser R, Westermann B. Optical tracking of a microscope for image-guided intranasal sinus surgery. *The Annals of otology, rhinology, and laryngology.* Jan 1999;108(1):54-62.
177. Zheng G, Caversaccio M, Bachler R, Langlotz F, Nolte LP, Hausler R. Frameless optical computer-aided tracking of a microscope for otorhinology and skull base surgery. *Archives of otolaryngology--head & neck surgery.* Oct 2001;127(10):1233-1238.
178. Kral F, Puschban EJ, Riechelmann H, Pedross F, Freysinger W. Optical and electromagnetic tracking for navigated surgery of the sinuses and frontal skull base. *Rhinology.* Aug 2011;49(3):364-368.
179. Kuo TH, Wong TY, Wu TC, Fang JJ. Optical tracking-based model surgery for orthognathic surgical planning using a quantifying evaluation method. *IEEE Trans Inf Technol Biomed.* Nov 2012;16(6):1193-1199.
180. Li B, Zhang L, Sun H, Shen SG, Wang X. A new method of surgical navigation for orthognathic surgery: optical tracking guided free-hand repositioning of the maxillomandibular complex. *The Journal of craniofacial surgery.* Mar 2014;25(2):406-411.
181. Cartellieri M, Vorbeck F. Endoscopic Sinus Surgery Using Intraoperative Computed Tomography Imaging for Updating a Three-Dimensional Navigation System. *The Laryngoscope.* 2000;110(2):292-292.
182. Navab N, Wiesner S, Benhimane S, Euler E, Heining SM. Visual servoing for intraoperative positioning and repositioning of mobile C-arms. *Med Image Comput Comput Assist Interv.* 2006;9(Pt 1):551-560.
183. P D, L W, O K, J T, S-M H, N N. Intraoperative positioning of mobile C-arms using artificial fluoroscopy. Paper presented at: Proc SPIE2010.
184. Navab N, Heining SM, Traub J. Camera augmented mobile C-arm (CAMC): calibration, accuracy study, and clinical applications. *IEEE Trans Med Imaging.* Jul 2010;29(7):1412-1423.



185. Reaungamornrat S, Otake Y, Uneri A, et al. An on-board surgical tracking and video augmentation system for C-arm image guidance. *Int J Comput Assist Radiol Surg.* Sep 2012;7(5):647-665.
186. Staub D, Murphy MJ. A digitally reconstructed radiograph algorithm calculated from first principles. *Med Phys.* Jan 2013;40(1):011902.
187. Moore CS, Avery G, Balcam S, et al. Use of a digitally reconstructed radiograph-based computer simulation for the optimisation of chest radiographic techniques for computed radiography imaging systems. *Br J Radiol.* Sep 2012;85(1017):e630-639.
188. Li X, Yang J, Zhu Y. Digitally reconstructed radiograph generation by an adaptive Monte Carlo method. *Phys Med Biol.* Jun 7 2006;51(11):2745-2752.
189. Freund E, Heinze F, Rossmann J. Direction dependent projection fields for the fast DRR generation for medical 2D/3D registration. *Conf Proc IEEE Eng Med Biol Soc.* 2004;3:1751-1754.
190. Birkfellner W, Seemann R, Figl M, et al. Fast DRR generation for 2D/3D registration. *Med Image Comput Comput Assist Interv.* 2005;8(Pt 2):960-967.
191. Chintalapani G, Jain A, Taylor RH. Statistical Characterization of C-arm Distortion with Application to Intra-Operative Distortion Correction. Paper presented at: SPIE Medical Imaging 2007: Image-Guided Procedures, Robotic Interventions, and Modeling 2007.
192. Brost A, Strobel N, Yatziv L, et al. Geometric Accuracy of 3-D X-Ray Image-Based Localization from Two C-Arm Views. Workshop on Geometric Accuracy In Image Guided Interventions - Medical Image Computing and Computer Assisted Interventions (MICCAI); 2009; London, UK.
193. Brost A, Strobel N, Yatziv L, et al. Accuracy of X-Ray Image-Based 3D Localization from Two C-Arm Views: A Comparison Between an Ideal System and a Real Device. *SPIE Medical Imaging.* 2009.
194. Jin J-Y, Ryu S, Faber K, et al. 2D/3D Image fusion for accurate target localization and evaluation of a mask based stereotactic system in fractionated stereotactic radiotherapy of cranial lesions. *Medical physics.* 2006;33(12):4557-4566.
195. Kowatsch M, Winkler P, Zurl B, Konrad T, Schrottner J. [Analysis of image quality and dose calculation accuracy in cone beam CT acquisitions with limited projection data (half scan, half fan) with regard to usability for adaptive radiation therapy treatment planning]. *Zeitschrift fur medizinische Physik.* 2011;21(1):11-18.
196. Yang D, Ning R. FDK Half-Scan with a Heuristic Weighting Scheme on a Flat Panel Detector-Based Cone Beam CT (FDKHSCW). *Int J Biomed Imaging.* 2006;2006:83983.
197. Lee SW, Wang G. A Grangeat-type half-scan algorithm for cone-beam CT. *Med Phys.* Apr 2003;30(4):689-700.
198. Wang G, Liu Y, Lin TH, Cheng PC. Half-scan cone-beam x-ray microtomography formula. *Scanning.* Jul-Aug 1994;16(4):216-220.
199. Zhang J, Wu QJ, Godfrey DJ, Fatunase T, Marks LB, Yin FF. Comparing digital tomosynthesis to cone-beam CT for position verification in patients undergoing partial breast irradiation. *Int J Radiat Oncol Biol Phys.* Mar 1 2009;73(3):952-957.
200. Gong X, Glick SJ, Liu B, Vedula AA, Thacker S. A computer simulation study comparing lesion detection accuracy with digital mammography, breast tomosynthesis, and cone-beam CT breast imaging. *Med Phys.* Apr 2006;33(4):1041-1052.
201. Lux MM, Marshall M, Erturk E, Joseph JV. Ergonomic evaluation and guidelines for use of the daVinci Robot system. *J Endourol.* Mar 2010;24(3):371-375.
202. Pick DL, Lee DI, Skarecky DW, Ahlering TE. Anatomic guide for port placement for daVinci robotic radical prostatectomy. *J Endourol.* Aug 2004;18(6):572-575.
203. Descovich M, Morin O, Aubry JF, et al. Characteristics of megavoltage cone-beam digital tomosynthesis. *Med Phys.* Apr 2008;35(4):1310-1316.
204. Karellas A, Lo JY, Orton CG. Point/Counterpoint. Cone beam x-ray CT will be superior to digital x-ray tomosynthesis in imaging the breast and delineating cancer. *Med Phys.* Feb 2008;35(2):409-411.
205. Pang G, Bani-Hashemi A, Au P, et al. Megavoltage cone beam digital tomosynthesis (MV-CBDT) for image-guided radiotherapy: a clinical investigational system. *Phys Med Biol.* Feb 21 2008;53(4):999-1013.

206. Park JC, Park SH, Kim JH, et al. Four-dimensional cone-beam computed tomography and digital tomosynthesis reconstructions using respiratory signals extracted from transcutaneously inserted metal markers for liver SBRT. *Med Phys.* Feb 2011;38(2):1028-1036.
207. Stayman JW, Dang H, Ding Y, Siewerdsen JH. PIRPLE: a penalized-likelihood framework for incorporation of prior images in CT reconstruction. *Phys Med Biol.* Nov 7 2013;58(21):7563-7582.
208. Ellingsen LM, Chintalapani G, Taylor RH, Prince JL. Robust deformable image registration using prior shape information for atlas to patient registration. *Comput Med Imaging Graph.* Jan 2010;34(1):79-90.
209. Otake Y, Schafer S, Stayman JW, et al. Automatic Localization of target vertebrae in spine surgery using fast CT-to-fluoroscopy (3D-2D) image registration. *SPIE Medical Imaging 2012: Image-Guided Procedures, Robotic Interventions, and Modeling.* 2012;8316:8316-8310N.
210. Uneri A, Otake Y, Wang AS, et al. 3D-2D registration for surgical guidance: effect of projection view angles on registration accuracy. *Phys Med Biol.* Dec 19 2013;59(2):271-287.
211. Jain A, Kon R, Zhou Y, Fichtinger G. C-arm calibration--is it really necessary? *Med Image Comput Comput Assist Interv.* 2005;8(Pt 1):639-646.
212. Kapoor A, Li M, Taylor RH. Spatial motion constraints for robot assisted suturing using virtual fixtures. *Med Image Comput Comput Assist Interv.* 2005;8(Pt 2):89-96.
213. Ali MR, Loggins JP, Fuller WD, et al. 3-D telestration: a teaching tool for robotic surgery. *Journal of laparoendoscopic & advanced surgical techniques. Part A.* Feb 2008;18(1):107-112.
214. Milgram P, Takemura H, Utsumi A, Kishino F. Augmented Reality: A class of displays on the reality-virtuality continuum Paper presented at: SPIE Telemanipulator and Telepresence Technologies1994.
215. Roberts DW, Strohbehn JW, Hatch JF, Murray W, Kettenberger H. A frameless stereotaxic integration of computerized tomographic imaging and the operating microscope. *J Neurosurg.* Oct 1986;65(4):545-549.
216. Friets EM, Strohbehn JW, Hatch JF, Roberts DW. A frameless stereotaxic operating microscope for neurosurgery. *IEEE Trans Biomed Eng.* Jun 1989;36(6):608-617.
217. Berger JW, Shin DS. Computer-vision-enabled augmented reality fundus biomicroscopy. *Ophthalmology.* Oct 1999;106(10):1935-1941.
218. Birkfellner W, Figl M, Huber K, et al. A head-mounted operating binocular for augmented reality visualization in medicine--design and initial evaluation. *IEEE Trans Med Imaging.* Aug 2002;21(8):991-997.
219. Bajura M, Fuchs HE, Ohbuchi R. Merging virtual objects with the real world: seeing ultrasound imagery within the patient. Paper presented at: Proceedings of the 19th annual conference on Computer graphics and interactive techniques1992.
220. Bichlmeier C, Heining SM, Feuerstein M, Navab N. The virtual mirror: a new interaction paradigm for augmented reality environments. *IEEE Trans Med Imaging.* Sep 2009;28(9):1498-1510.
221. Bichlmeier C, Heining SM, Rustae M, Navab N. Virtually extended surgical drilling device: virtual mirror for navigated spine surgery. *Med Image Comput Comput Assist Interv.* 2007;10(Pt 1):434-441.
222. Muller M, Rassweiler MC, Klein J, et al. Mobile augmented reality for computer-assisted percutaneous nephrolithotomy. *Int J Comput Assist Radiol Surg.* Jul 2013;8(4):663-675.
223. Wagner D, Schmalstieg D. Making augmented reality practical on mobile phones, part 2. *IEEE computer graphics and applications.* Jul-Aug 2009;29(4):6-9.
224. Wagner D, Schmalstieg D. Making augmented reality practical on mobile phones, part 1. *IEEE computer graphics and applications.* May-Jun 2009;29(3):12-15.
225. Glauser W. Doctors among early adopters of Google Glass. *CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne.* Nov 5 2013;185(16):1385.
226. Edwards PJ, King AP, Hawkes DJ, et al. Stereo augmented reality in the surgical microscope. *Stud Health Technol Inform.* 1999;62:102-108.
227. Pratt P, Edwards E, Arora A, Tolley N, Darzi AW, Yang G-Z. Image-guided transoral robotic surgery for the treatment of oropharyngeal cancer. Paper presented at: Hamlyn Symposium2012; London.

228. Falk V, Mourgues F, Vieville T, et al. Augmented reality for intraoperative guidance in endoscopic coronary artery bypass grafting. *Surgical technology international*. 2005;14:231-235.
229. Klein T, Hansson M, Navab N. Modeling of multi-view 3D freehand radio frequency ultrasound. *Med Image Comput Comput Assist Interv*. 2012;15(Pt 1):422-429.
230. Daly MJ, Siewerdsen JH, Moseley DJ, Jaffray DA, Irish JC. Intraoperative cone-beam CT for guidance of head and neck surgery: Assessment of dose and image quality using a C-arm prototype. *Med Phys*. Oct 2006;33(10):3767-3780.
231. Whitman J, Fronheiser MP, Ivancevich NM, Smith SW. Autonomous surgical robotics using 3-D ultrasound guidance: feasibility study. *Ultrason Imaging*. Oct 2007;29(4):213-219.
232. Foster W, Beaulieu L, Harel F, Martin AG, Vigneault E. The impact of 3D image guided prostate brachytherapy on therapeutic ratio: the Quebec University Hospital experience. *Cancer radiotherapie : journal de la Societe francaise de radiotherapie oncologique*. Dec 2007;11(8):452-460.
233. Nguyen PL, Chen MH, D'Amico AV, et al. Magnetic resonance image-guided salvage brachytherapy after radiation in select men who initially presented with favorable-risk prostate cancer: a prospective phase 2 study. *Cancer*. Oct 1 2007;110(7):1485-1492.
234. Krieger A, Csoma C, Iordachital, II, et al. Design and preliminary accuracy studies of an MRI-guided transrectal prostate intervention system. *Med Image Comput Comput Assist Interv*. 2007;10(Pt 2):59-67.
235. Mozer P, Troccaz J, Stoianovici D. Urologic robots and future directions. *Current opinion in urology*. Jan 2009;19(1):114-119.
236. Chang PL, Stoyanov D, Davison AJ, Edwards PE. Real-time dense stereo reconstruction using convex optimisation with a cost-volume for image-guided robotic surgery. *Med Image Comput Comput Assist Interv*. 2013;16(Pt 1):42-49.
237. Stoyanov D, Scarzanella M, Pratt P, Yang G-Z. Real-Time Stereo Reconstruction in Robotically Assisted Minimally Invasive Surgery. In: Jiang T, Navab N, Pluim J, Viergever M, eds. *Medical Image Computing and Computer-Assisted Intervention, MICCAI 2010*. Vol 6361: Springer Berlin / Heidelberg; 2010:275-282.
238. Khamene A, Wacker F, Vogt S, et al. An Augmented Reality system for MRI-guided needle biopsies. *Stud Health Technol Inform*. 2003;94:151-157.
239. Wendt M, Sauer F, Khamene A, Bascle B, Vogt S, Wacker FK. [A head-mounted display system for augmented reality: initial evaluation for interventional MRI]. *RoFo : Fortschritte auf dem Gebiete der Rontgenstrahlen und der Nuklearmedizin*. Mar 2003;175(3):418-421.
240. Stetten G, Chib V. Overlaying ultrasound images on direct vision. *Int. J. Ultrasound in Medicine*. 2001;20:235-240.
241. Fichtinger G, Deguet A, Masamune K, et al. Image overlay guidance for needle insertion in CT scanner. *IEEE Trans Biomed Eng*. Aug 2005;52(8):1415-1424.
242. Fischer GS, Deguet A, Csoma C, et al. MRI image overlay: application to arthrography needle insertion. *Comput Aided Surg*. Jan 2007;12(1):2-14.
243. Xu C, Pham D, Prince J. Medical Image Segmentation Using Deformable Models. *Handbook of Medical Imaging*. Vol 2: Medical Image Processing and Analysis: SPIE Press May; 2000.
244. Yushkevich PA, Piven J, Hazlett HC, et al. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *NeuroImage*. 2006;31(3):1116-1128.
245. P. A. Yushkevich JP, H Cody, S. Ho, J.C. Gee and G. Gerig. User-Guided Level Set Segmentation of Anatomical Structures with ITK-SNAP. *Insight Journal*. 2005;1.
246. Fedorov A, Beichel R, Kalpathy-Cramer J, et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magn Reson Imaging*. Nov 2012;30(9):1323-1341.
247. Sielhorst T, Feuerstein M, Navab N. Advanced Medical Displays: A Literature Review of Augmented Reality. *J Disp Technol*. 2008;4(4).
248. Mori K, Deguchi D, Akiyama K, et al. Hybrid bronchoscope tracking using a magnetic tracking sensor and image registration. *Med Image Comput Comput Assist Interv*. 2005;8(Pt 2):543-550.

249. Klein T, Traub J, Hautmann H, Ahmadian A, Navab N. Fiducial-free registration procedure for navigated bronchoscopy. *Med Image Comput Comput Assist Interv.* 2007;10(Pt 1):475-482.
250. Bricault I, Ferretti G, Cinquin P. Registration of real and CT-derived virtual bronchoscopic images to assist transbronchial biopsy. *IEEE Trans Med Imaging.* Oct 1998;17(5):703-714.
251. Bricault I, Ferretti G, Cinquin P. Computer-assisted bronchoscopy: aims and research perspectives. *Journal of image guided surgery.* 1995;1(4):217-225.
252. Mori K, Deguchi D, Sugiyama J, et al. Tracking of a bronchoscope using epipolar geometry analysis and intensity-based image registration of real and virtual endoscopic images. *Med Image Anal.* Sep 2002;6(3):321-336.
253. Wann JP, Rushton S, Mon-Williams M. Natural problems for stereoscopic depth perception in virtual environments. *Vision Res.* Oct 1995;35(19):2731-2736.
254. Sielhorst T, Bichlmeier C, Heining SM, Navab N. Depth perception--a major issue in medical AR: evaluation study by twenty surgeons. *Med Image Comput Comput Assist Interv.* 2006;9(Pt 1):364-372.
255. Bichlmeier C, Wimpe F, Heining SM, Navab N. Contextual Anatomic Mimesis Hybrid In-Situ Visualization Method for Improving Multi-Sensory Depth Perception in Medical Augmented Reality. Paper presented at: IEEE and ACM International Symposium on Mixed and Augmented Reality 2007.
256. Stayman JW, Otake Y, Prince JL, Khanna AJ, Siewerdsen JH. Model-based tomographic reconstruction of objects containing known components. *IEEE transactions on medical imaging.* Oct 2012;31(10):1837-1848.
257. Mirota D, Wang H, Taylor R, Ishii M, Hager G. Toward Video-Based Navigation for Endoscopic Endonasal Skull Base Surgery. MICCAI; 2009; Berlin, Germany.
258. Jannin P, Korb W. *Image-Guided Interventions.* Springer US; 2008.
259. Cohen D, Mayer E, Chen D, Anstee A, Vale J, Yang GZ. Augmented reality image guidance in minimally invasive prostatectomy. *Prostate Cancer Imaging.* 2010;6367:101-110.
260. Viant WJ. The development of an evaluation framework for the quantitative assessment of computer-assisted surgery and augmented reality accuracy performance. *Stud Health Technol Inform.* 2001;81:534-540.
261. Jannin P, Fitzpatrick JM, Hawkes DJ, Pennec X, Shahidi R, Vannier MW. Validation of medical image processing in image-guided therapy. *IEEE Trans Med Imaging.* Dec 2002;21(12):1445-1449.
262. Siewerdsen JH, Moseley DJ, Burch S, et al. Volume CT with a flat-panel detector on a mobile, isocentric C-arm: pre-clinical investigation in guidance of minimally invasive surgery. *Med Phys.* Jan 2005;32(1):241-254.
263. Siewerdsen JH, Daly MJ, Bachar G, et al. Multimode C-arm fluoroscopy, tomosynthesis, and cone-beam CT for image-guided interventions: from proof of principle to patient protocols. *Proc. of SPIE Medical Imaging.* 2007:65101A-65101A.
264. Siewerdsen JH. Cone-Beam CT with a Flat-Panel Detector: From Image Science to Image-Guided Surgery. *Nucl Instrum Methods Phys Res A.* Aug 21 2011;648(S1):S241-S250.
265. Liu WP, Richmon JD, Azizian M, Sorger JM, Taylor RH. Intraoperative Cone Beam CT Guidance for Transoral Robotic Surgery. Paper presented at: Information Processing in Computer Assisted Interventions (IPCAI) 2014; Fukuoka, Japan.
266. Liu WP, Reaungamornrat S, A. D, et al. Toward Intraoperative Image-Guided Transoral Robotic Surgery. *Robotic Surgery.* 2013;7(2):217-225.
267. Reiter A, Allen PK, Zhao T. Feature classification for tracking articulated surgical tools. *Med Image Comput Comput Assist Interv.* 2012;15(Pt 2):592-600.
268. Rieger A, Blum T, Navab N, Friess H, Martignoni ME. [Augmented reality: merge of reality and virtuality in medicine]. *Deutsche medizinische Wochenschrift.* Nov 2011;136(47):2427-2433.
269. Liu W, Azizian M, Sorger JM, et al. A Pilot Study of Augmented Reality from Intraoperative CBCT for Image-Guided Thoracic Robotic Surgery. The Hamlyn Symposium on Medical Robotics; 2014; London, U. K.

270. Katsuyama Y, Yamazaki N, Kobayashi Y, Hoshi T, Miyashita T. A study on estimation of the deformation behavior in the collapse process of lung. *Conf Proc IEEE Eng Med Biol Soc.* 2012;2012:2817-2822.
271. Schulze F, Buhler K, Neubauer A, Kanitsar A, Holton L, Wolfsberger S. Intra-operative virtual endoscopy for image guided endonasal transsphenoidal pituitary surgery. *Int J Comput Assist Radiol Surg.* Mar 2010;5(2):143-154.
272. Lapeer R, Chen MS, Gonzalez G, Linney A, Alusi G. Image-enhanced surgical navigation for endoscopic sinus surgery: evaluating calibration, registration and tracking. *The international journal of medical robotics + computer assisted surgery : MRCAS.* Mar 2008;4(1):32-45.
273. Kawamata T, Iseki H, Shibasaki T, Hori T. Endoscopic augmented reality navigation system for endonasal transsphenoidal surgery to treat pituitary tumors: technical note. *Neurosurgery.* Jun 2002;50(6):1393-1397.
274. Siewerdsen JH, Daly MJ, Chan H, et al. High performance intraoperative cone-beam CT on a mobile C-arm: an integrated system for guidance of head and neck surgery. Paper presented at: SPIE Medical Imagin2009.
275. Fraser JF, Allen B, Anand VK, Schwartz TH. Three-dimensional neurostereoeendoscopy: subjective and objective comparison to 2D. *Minim Invasive Neurosurg.* Feb 2009;52(1):25-31.
276. Kassam AB, Prevedello DM, Carrau RL, et al. Endoscopic endonasal skull base surgery: analysis of complications in the authors' initial 800 patients. *J Neurosurg.* Jun 2011;114(6):1544-1568.
277. Liu WP, Azizian M, Sorger J, et al. Cadaveric feasibility study of da Vinci Si-assisted cochlear implant with augmented visual navigation for otologic surgery. *JAMA otolaryngology-- head & neck surgery.* Mar 2014;140(3):208-214.
278. Wanna GB, Balachandran R, Majdani O, Mitchell J, Labadie RF. Percutaneous access to the petrous apex in vitro using customized micro-stereotactic frames based on image-guided surgical technology. *Acta oto-laryngologica.* Aug 25 2009;1-6.
279. Bell B, Stieger C, Gerber N, et al. A self-developed and constructed robot for minimally invasive cochlear implantation. *Acta oto-laryngologica.* Apr 2012;132(4):355-360.
280. Baron S, Eilers H, Munske B, et al. Percutaneous inner-ear access via an image-guided industrial robot system. *Proceedings of the Institution of Mechanical Engineers. Part H, Journal of engineering in medicine.* 2010;224(5):633-649.
281. Danilchenko A, Balachandran R, Toennies JL, et al. Robotic mastoidectomy. *Otol Neurotol.* Jan 2011;32(1):11-16.
282. Kasahara Y, Kawana H, Usuda S, Ohnishi K. Telerobotic-assisted bone-drilling system using bilateral control with feed operation scaling and cutting force scaling. *The international journal of medical robotics + computer assisted surgery : MRCAS.* Jun 2012;8(2):221-229.
283. Federspil PA, Geisthoff UW, Henrich D, Plinkert PK. Development of the first force-controlled robot for otoneurosurgery. *Laryngoscope.* Mar 2003;113(3):465-471.
284. Labadie RF, Balachandran R, Mitchell JE, et al. Clinical validation study of percutaneous cochlear access using patient-customized microstereotactic frames. *Otol Neurotol.* Jan 2010;31(1):94-99.
285. Labadie RF, Noble JH, Dawant BM, Balachandran R, Majdani O, Fitzpatrick JM. Clinical validation of percutaneous cochlear implant surgery: initial report. *Laryngoscope.* Jun 2008;118(6):1031-1039.
286. Balachandran R, Mitchell JE, Blachon G, et al. Percutaneous cochlear implant drilling via customized frames: an in vitro study. *Otolaryngol Head Neck Surg.* Mar 2010;142(3):421-426.
287. Zhang J, Wei W, Ding J, Roland JT, Jr., Manolidis S, Simaan N. Inroads toward robot-assisted cochlear implant surgery using steerable electrode arrays. *Otol Neurotol.* Oct 2010;31(8):1199-1206.
288. Roland PS, Wright CG, Isaacson B. Cochlear implant electrode insertion: the round window revisited. *The Laryngoscope.* Aug 2007;117(8):1397-1402.

289. Brown KD, Connell SS, Balkany TJ, Eshraghi AE, Telischi FF, Angeli SA. Incidence and indications for revision cochlear implant surgery in adults and children. *Laryngoscope*. Jan 2009;119(1):152-157.
290. Chung D, Kim AH, Parisier S, et al. Revision cochlear implant surgery in patients with suspected soft failures. *Otol Neurotol*. Oct;31(8):1194-1198.
291. Rau TS, Hussong A, Leinung M, Lenarz T, Majdani O. Automated insertion of preformed cochlear implant electrodes: evaluation of curling behaviour and insertion forces on an artificial cochlear model. *International journal of computer assisted radiology and surgery*. Mar 2010;5(2):173-181.
292. Majdani O, Schurzig D, Hussong A, et al. Force measurement of insertion of cochlear implant electrode arrays in vitro: comparison of surgeon to automated insertion tool. *Acta oto-laryngologica*. 2010;130(1):31-36.
293. Majdani O, Rau TS, Baron S, et al. A robot-guided minimally invasive approach for cochlear implant surgery: preliminary results of a temporal bone study. *International journal of computer assisted radiology and surgery*. Sep 2009;4(5):475-486.
294. Coulson CJ, Taylor RP, Reid AP, Griffiths MV, Proops DW, Brett PN. An autonomous surgical robot for drilling a cochleostomy: preliminary porcine trial. *Clinical otolaryngology : official journal of ENT-UK ; official journal of Netherlands Society for Oto-Rhino-Laryngology & Cervico-Facial Surgery*. Aug 2008;33(4):343-347.
295. Brett PN, Taylor RP, Proops D, Coulson C, Reid A, Griffiths MV. A surgical robot for cochleostomy. *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*. 2007;2007:1229-1232.
296. Metson RB, Cosenza MJ, Cunningham MJ, Randolph GW. Physician experience with an optical image guidance system for sinus surgery. *The Laryngoscope*. Jun 2000;110(6):972-976.
297. Thom JJ, Carlson ML, Olson MD, et al. The prevalence and clinical course of facial nerve paresis following cochlear implant surgery. *Laryngoscope*. Apr;123(4):1000-1004.
298. Kersten-Oertel M, Jannin P, Collins DL. The state of the art of visualization in mixed reality image guided surgery. *Comput Med Imaging Graph*. Mar 2013;37(2):98-112.
299. Zender J, Thell C. Developing a successful robotic surgery program in a rural hospital. *AORN journal*. Jul 2010;92(1):72-83; quiz 84-76.
300. Inc. IS. *Investor Presentation Q1 2014*. 2014.
301. Kazanzides P. Robot assisted surgery: the ROBODOC experience. Paper presented at: 30th International Symposium on Robotics1999; Tokyo.
302. Linte CA, Davenport KP, Cleary K, et al. Augmented environments for minimally invasive therapy: implementation barriers from technology to practice. *Stud Health Technol Inform*. 2012;173:263-269.
303. English J, Chang CY, Tardella N, Hu J. A vision-based surgical tool tracking approach for untethered surgery simulation and training. *Stud Health Technol Inform*. 2005;111:126-132.
304. Cleary K, Peters TM. Image-guided interventions: technology review and clinical applications. *Annu Rev Biomed Eng*. Aug 15 2010;12:119-142.
305. Birkfellner W, Watzinger F, Wanschitz F, Ewers R, Bergmann H. Calibration of tracking systems in a surgical environment. *IEEE Trans Med Imaging*. Oct 1998;17(5):737-742.
306. Hummel J, Figl M, Birkfellner W, et al. Evaluation of a new electromagnetic tracking system using a standardized assessment protocol. *Phys Med Biol*. May 21 2006;51(10):N205-210.
307. Kwartowitz DM, Rettmann ME, Holmes DRI, Robb RA. A novel technique for analysis of accuracy of magnetic tracking systems used in image-guided surgery. Paper presented at: SPIE Medical Imaging Symposium: Visualization, Image-guided Procedures and Modeling,2010.
308. Xing F, Woo J, Murano EZ, Lee J, Stone M, Prince JL. 3D tongue motion from tagged and cine MR images. *Med Image Comput Comput Assist Interv*. 2013;16(Pt 3):41-48.
309. Xing F, Murano EZ, Lee J, Woo J, Stone M, Prince JL. Mri Analysis of 3d Normal and Post-Glossectomy Tongue Motion in Speech. *Proceedings / IEEE International Symposium on*

- Biomedical Imaging: from nano to macro. IEEE International Symposium on Biomedical Imaging.* Dec 31 2013;2013:816-819.
310. Stone M, Langguth JM, Woo J, Chen H, Prince JL. Tongue Motion Patterns in Glossectomy and Normal Speakers: A Principal Components Analysis. *Journal of speech, language, and hearing research : JSLHR.* Sep 1 2013.
  311. Zhen X, Yan H, Zhou L, Jia X, Jiang SB. Deformable image registration of CT and truncated cone-beam CT for adaptive radiation therapy. *Phys Med Biol.* Nov 21 2013;58(22):7979-7993.
  312. Lou Y, Niu T, Jia X, Vela PA, Zhu L, Tannenbaum AR. Joint CT/CBCT deformable registration and CBCT enhancement for cancer radiotherapy. *Med Image Anal.* Apr 2013;17(3):387-400.
  313. Uneri A, Nithiananthan S, Schafer S, et al. Deformable registration of the inflated and deflated lung in cone-beam CT-guided thoracic surgery: initial investigation of a combined model- and image-driven approach. *Med Phys.* Jan 2013;40(1):017501.
  314. Sotiras A, Davatzikos C, Paragios N. Deformable medical image registration: a survey. *IEEE Trans Med Imaging.* Jul 2013;32(7):1153-1190.
  315. Schrader P. [Technique evaluation for orthopedic use of Robodoc]. *Zeitschrift fur Orthopadie und ihre Grenzgebiete.* May-Jun 2005;143(3):329-336.
  316. Hananouchi T, Nakamura N, Kakimoto A, Yohsikawa H, Sugano N. CT-based planning of a single-radius femoral component in total knee arthroplasty using the ROBODOC system. *Comput Aided Surg.* Jan 2008;13(1):23-29.

## Vita

Wen Pei Liu was made in Taiwan and grew up in San Diego, California. A high school salutatorian lettering in varsity basketball, she continued her amateur pursuit of sports at Johns Hopkins where her coed basketball team claimed the intramural championships three years running (2010-2013). She received a Bachelor's Degree in Computer Science with a Minor in Mechanical Engineering from the University of California, Berkeley (2004) and a Masters in Electrical Engineering from Stanford University (2008). Her studies at Stanford was supported through a continuing academic program at the NASA Ames Research Center, where she was employed as a research scientist at the Advanced Supercomputing Division (2004-2006) and a software engineer at the Intelligent Systems Division (2006-2009). For Project Columbia, which received

an honorable mention for NASA's Software of the Year Award (2011), Wen supervised the Scheduling Module development, leading her team to become the first group in history to retire parts of the original Scheduling Module in the Training Management System used by the Mission Operations Directorate. During that time, she cultivated an interest in medical robotics during a rotation in the BioRobotics Lab at Stanford and during an internship at Hansen Medical, Inc. (2009).

For her PhD at the Johns Hopkins University, Wen's inaugural projects were supported by the Johns Hopkins University Applied Physics Laboratory and the BioMedical Engineering department at Johns Hopkins School of Medicine. These two projects looked at image guidance for bone augmentation to treat osteolysis in orthopedic surgery, and efforts in clinical engineering to ready an IRB study by using image guidance through video augmentation for endoscopic skull base surgery, respectively. Her thesis work on developing an innovative intraoperative image-guidance system for robotic surgery using cone-beam computed tomography was supported through a fellowship through Intuitive Surgical, Inc.

In 2014, Wen joined Siemens Corporation, Corporate Technology as a Research Scientist and continues to pursue her research interests in image-guided medical robotics, interventional imaging, clinical engineering, and perceptual computing.



# Curriculum Vitae

Name: Wen Pei Liu

Date of Birth: 27 January 1982

Location of Birth: Taipei, Taiwan

Nationality: United States of America

Johns Hopkins University (LCSR/CISST-ERC)

3400 North Charles Street; Hackerman Hall rm. 136; Baltimore, MD 21218; USA

Website: <https://www.lcsr.jhu.edu/User:Wliu25>

Email Address: [wen.p.liu@gmail.com](mailto:wen.p.liu@gmail.com)

Mobile Phone: 510.717.3327

## Educational History

2014 Doctor of Philosophy

Johns Hopkins University, Baltimore, MD

Program in Computer Science

Advisor: Russell H. Taylor, PhD

2011 Master of Science

Johns Hopkins University, Baltimore, MD

Program in Computer Science

2011 Master of Science

Stanford University, Stanford, CA

Program in Electrical Engineering

Certificate, Computer Networks

2011 Bachelor of Arts

University of California, Berkeley, CA

Program in Computer Science

Minor, Mechanical Engineering

## **Professional Experience**

June 2014 - Present Research Scientist

Imaging and Computer Vision

Siemens Corporation, Corporate Technology, Princeton, NJ

*Research in algorithms and devices in registration and visualization of multimodal interventional imaging and image guidance.*

January 2012 - June 2014 Research Assistant

Laboratory for Computational Sensing and Robotics

Johns Hopkins University, Baltimore, MD

*My doctoral thesis integrates image guidance through cone-beam computed tomography (CBCT) with minimally invasive surgical systems.*

June 2012 - June 2014 Medical Research Intern

Intuitive Surgical, Inc. Sunnyvale, CA

*Developing workflow and software for intraoperative image guidance through cone-beam computed tomography (CBCT) with minimally invasive, da Vinci robotic surgical systems.*

September 2011 - December 2011 Teaching Assistant

Course: Algorithms for Sensor-based Robotics, Computer Science

Johns Hopkins University, Baltimore, MD

*Assisted lectures, create solutions and grading matrix for programming/written assignments, projects and exams for course to introduce the fundamental topics of basic sensing and estimation techniques, geometric model representations, and motion planning algorithms in mobile robotics.*

January 2011 - August 2011 Research Assistant

Imaging for Surgery, Therapy and Radiology (ISTAR) Laboratory

Johns Hopkins University, Baltimore, MD

*Researched in image guided surgery by augmentation of monocular endoscopic video with CT/MR in minimally invasive skull base surgery.*

January 2011 - August 2011 Research Assistant

Applied Physics Laboratory

Johns Hopkins University, Baltimore, MD

*Research development for a minimally invasive percutaneous femoroplasty surgical system, sponsored by the National Institute of Health.*

January 2007 - August 2009 Research Assistant

BioRobotics Laboratory

Stanford University, Stanford, CA

*Developed a visuohaptic temporal bone surgical simulator in collaboration with  
Stanford University Medical Center, Otolaryngology Department*

July 2006 - August 2009 Software Development Lead

NASA, Intelligent Systems Division

NASA Ames Research Center, Moffett Field, CA

*Operational and development responsibilities for advanced spaceflight training  
management system in support of NASA mission operations, including the Space  
Shuttle, International Space Station and Constellation project.*

June 2008 - September 2008 Software Engineer Intern

Hansen Medical, Inc.

Mountain View, CA

*Proposed and implemented R-Function logical combinations of implicit functions  
using known primitive implicit building blocks to improve the haptic-rendering of  
various catheter control workspaces for the Sensei Robotic Catheter System*

April 2008 - June 2008 Course Assistant

Course: Experimental Haptics, Computer Science

Stanford University, Stanford, CA

*Assisted lectures, grading for course which introduces the essential topics of haptics programming and touch feedback in virtual reality.*

July 2004 - July 2006 Computer Scientist

NASA Advanced Supercomputing Division

NASA Ames Research Center, Moffett Field, CA

*Member of a team responsible for a new security enclave for Columbia, a 10,240 Itanium II supercomputer.*

January 2003 - May 2004 Research Assistant

Robotics and Intelligent Machines Laboratory

University of California Berkley, Berkeley, CA

*Automated the assembly of mechatronic, biomimetric devices through the process, "Desktop Rapid Prototyping Millirobots"*

September 2002 - January 2003 Research Assistant

Mechanical Engineering Laboratory

University of California Berkley, Berkeley, CA

*Integrated material science properties with applications of electrical and physical science by exploring a shape memory alloy power, "Nanomuscles", to design a mechanical fish.*

## **Publications**

- **W. P. Liu**, S. Reagamornrat, J. M. Sorger, J. H. Siewerdsen, R. H. Taylor, J. D. Richmon, "Intraoperative Image-Guided Transoral Robotic Surgery: Pre-Clinical

Studies", *Int Journal Medical Robotics Computer-assisted Surgery (IJMRCAS)*, 2014."

- **W. P. Liu**, M. Azizian, J. M. Sorger, B. Mungo, O. Wagner, D. Molena, R. H. Taylor, "A Pilot Study of Augmented Reality from Intraoperative CBCT for Image-Guided Thoracic Robotic Surgery", *Proc. Hamlyn Symposium on Medical Robotics, London, 2014.*"
- **W. P. Liu**, J. D. Richmon, M. Azizian, J. M. Sorger, R. H. Taylor, "Intraoperative Cone Beam CT Guidance for Transoral Robotic Surgery", *Information Processing in Computer Assisted Intervention (IPCAI) 2014.*
- **W. P. Liu**, M. Azizian, J. Sorger, R. H. Taylor, B. Reilly, K. Cleary, D. Preciado, "A Cadaveric Feasibility Study of da Vinci-Assisted Cochlear Implantation with Augmented Visual Navigation for Otologic Surgery", *Journal of the American Medical Association-Otolaryngology Head and Neck Surgery*, November 2013
- **W. P. Liu**, S. Reaugamornrat, A. Deguet, J. M. Sorger, J. H. Siewerdsen, J. Richmon, R. H. Taylor, "Toward Intraoperative Image-Guided Transoral Robotic Surgery", *Journal of Robotic Surgery*, July 2013.
- Reaugamornrat S, **Liu WP**, Wang AS, Otake Y, Nithiananthan S, Uneri A, Schafer S, Tryggestad E, Richmon J, Sorger JM, Siewerdsen JH, and Taylor RH, "Deformable registration for cone-beam CT-guided transoral robotic base of tongue surgery," *Phys. Med. Biol.* 2013.
- S. Reaugamornrat, **W. P. Liu**, S. Schafer, Y. Otake, S. Nithiananthan, A. Uneri, J. Richmon, J. Sorger, J. H. Siewerdsen, R. H. Taylor, "Deformable Registration for Cone-Beam CT Guidance of Robot-Assisted, Trans-Oral Base-of-Tongue Surgery" in *Society of Photographic Instrumentation Engineers (SPIE) Medical Imaging Conference*, 2013

- J. Richmon, **W. P. Liu**, R. H. Taylor, S. Reaugamornrat, J. Sorger, J. H. Siewerdsen, "Intraoperative Image-Guided Transoral Robotic Surgery: Pre-Clinical Studies", 20th International Federation of Oto-Rhino-Laryngological Societies (IFOS 2013)
- **W. P. Liu**, S. Reaugamornrat, A. Deguet, J. M. Sorger, J. H. Siewerdsen, J. Richmon, R. H. Taylor, "Toward Intraoperative Image-Guided TransOral Robotic Surgery", Proc. Hamlyn Symposium on Medical Robotics, pp44-45, London, 2012.
- **W. P. Liu**, D. Mirota, A. Uneri, Y. Otake, G. Hager, D. Reh, M. Ishii, G. Gallia, J. Siewerdsen, "A Clinical Pilot Study of a Modular Video-CT Augmentation System for Image-Guided Skull Base Surgery," in Society of Photographic Instrumentation Engineers (SPIE) Medical Imaging Conference, 2012.
- **W. P. Liu**, B.C. Lucas, K. Guerin, and E. Plaku, "Sensor and Sampling-based Motion Planning for Minimally Invasive Robotic Exploration of Osteolytic Lesions," in Intelligent Robots and Systems (IROS), IEEE/RSJ International Conference, 2011.
- **W. P. Liu**, M. Armand, Y. Otake, and R. H. Taylor, "Optimization of multi-image pose recovery of fluoroscope tracking (FTRAC) fiducial in an image-guided femoroplasty system," in Society of Photographic Instrumentation Engineers (SPIE) Medical Imaging Conference, 2011.
- **W. P. Liu**, C. Sewell, N. Blevins, K. Salisbury, K. Bodin, and N. Hjelte, "Representing Fluid with Smoothed Particle Hydrodynamics in a Cranial Base Simulator," in Medicine Meet Virtual Reality (MMVR), 2008.

## **Awards and Honors**

- NASA Software of the Year for Project STAR, Honorable Mention July 2011
- Johns Hopkins University, CS Howard and Jacqueline Chertkof Fellowship Sept 2009

- NASA Graduate Student Researcher's Program Fellowship 2009
- NASA Group Achievement Award for JSC MOD ITA with ARC April 16, 2009
- NASA Group Achievement Award for Project Columbia April 4, 2005
- University of California, Berkeley, Wooley Davis Scholarship 2003-2004
- University of California, Berkeley, J.W. Jones Scholarship 2000-2001
- University of California, Berkeley, Walter Hass Scholarship 2000-2001

### **Service and Leadership**

- Moveable Feast, Baltimore Community Organization Volunteer 2013
- Johns Hopkins University, Women's Club Basketball. (Treasurer) 2009-2014
- Co-ed Intramural Basketball, Johns Hopkins University (captain). 2009-2014
- Johns Hopkins University, Habitat for Humanity Volunteer 2010-2012
- Johns Hopkins University, Women of Whiting Outreach Officer. 2010-2011
- Johns Hopkins University, Crossover Program Volunteer. 2010-11
- Co-ed Intramural Basketball, Stanford University (captain). 2006-2008
- Asian American Donor Program Volunteer 2008
- Society of Women Engineers: Junior High School Outreach Officer 2001–2002
- Student Tutorial Endeavor Program (STEP) Volunteer 2001



